CrossFire Gmelin
A Tutorial

BEILSTEIN
INFORMATION

BEILSTEIN
INFORMATION

BEILSTEIN
INFORMATION

Distributed by:
Beilstein Informationssysteme GmbH
Varrentrappstrasse 40-42
D- 60486 Frankfurt
Germany
Tel: +49-69-7917-410
Fax: +49-69-7917-321

Help-Desk/CustomerService:
Beilstein Informationssysteme GmbH
Varrentrappstrasse 40-42
D- 60486 Frankfurt
Germany
Tel.: +49-69-7917-258
Fax.: +49-69-7917-321

Manual Version 2.3, July 1997, for Beilstein Commander Version 2.3 and CrossFire Server Version 3.x

Table of Contents

Introduction

1

1 Introduction

The Gmelin File under CrossFire can be thought of as a map of the world of inorganic and organometallic chemistry. This world is still largely unexplored, because it has not been possible in the past for any individual researcher to keep abreast of the vast amount of information published. Now for the first time you can quickly make your own inroads into charted and even uncharted territory. CrossFire makes this possible because its interactive nature and quick response encourage your creative curiosity. It's as simple as that.

The purpose of this short tutorial is to introduce you to the power of CrossFire, and to start you on the path which we hope will ultimately lead you to exciting new results in your own research. The knowledge and experience of generations of chemists is waiting ready to work for you. We know that once you start, the chemical literature will open up for you as never before. So why not get a taste of CrossFire now? You can always stop whenever you want to ...

The problem is: *you might never want to stop*. Be warned.

Best regards,

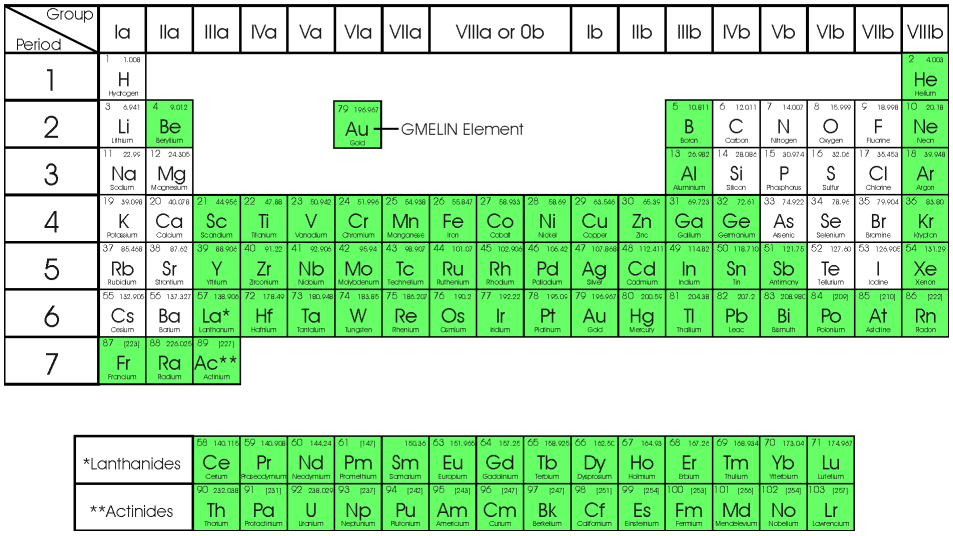
Your Beilstein Editorial Staff.

1.1 Contents of the Database

The Gmelin File covers inorganic and organometallic compounds. The following criteria are applied for the registration of a substance in the Gmelin database:

- All compounds containing no carbon
- All compounds containing at least one "Gmelin Element" (see periodic table below)
- Additionally: elemental carbon, phosgene, alloys, carbides and carbide oxides, carbonic acid and its thio- and seleno analogues, CO, CS, CO_2, CS_2, COS and multicomponent systems with a carbon component:

Periodic System of the Elements



The sources of the Gmelin file are as follows:

- Gmelin Handbook from 1772 to 1975
- Primary literature from 1975 - 1994

On the whole you will find in the Gmelin file:

- more than one million substance records
- about 900.000 reactions

- more than three million properties (within 800 different data fields)

1.2 The CrossFire Philosophy

You will use CrossFire to find information on some particular subject. The major challenge for any information retrieval system is to properly interpret your wishes in this respect. In CrossFire, you will first define the *Context* of your search, proceed to formulate your *Query* and then manipulate the *Display* of the results which are found. At each of these stages you will have a great amount of flexibility open to you, and it is not the intention of this tutorial to fully describe all the multitude of possibilities open to you at each point (that sort of documentation can be found in our User's Manual). Instead, we will try to paint the broader picture, and simply list some of the detailed aspects which you might like to follow up.

Your general *context* will usually be one (or a combination) of the two prime search objects:

- substances (and their specific properties including chemical reactions)
- documents (which deal with a specific topic)

When you set the context to one of these objects, the subsequent processing will deliver the results in the form of a *hitset* (or list) of objects of the type specified.

In setting the *query* you will generally draw chemical structure(s) and/or formulate factual queries. You can also combine these with previously obtained results from other queries.

In using the *display* you will have the option of examining your hitset of objects (i.e. documents or substances) and trimming these down to a more relevant subset if necessary, before looking at each in detail. At this point you will find the *hyperlinks* to be a fascinating means of following a trail at a tangent to your original query, changing the context "on the fly". At any time you can fire off new queries in either context.

This is the CrossFire philosophy in a nutshell.

PS: The tutorial was created using the Gmelin database as of 1997. As user of this text, you will maybe find minor discrepancies between your screen view and the figures used, since the database has been continuously updated since then. In particular, the tutorial was carried out using the operating system Microsoft Windows, on an IBM-compatible client. Users of the Macintosh version should find that the general description of the tutorial is equally applicable, since exactly the same philosophy was applied to their environment, with one not able exception: because of technical restrictions beyond our control, it was not possible to have a Child-window for structures in the Full Display window of the data on the Macintosh. This

means that the Mac user has the choice of displaying **either** structure graphics **or** text information to his individual hits, but not both simultaneously. There is of course a quick toggle between the two views with a single button click.

Searching a Substance by a Structure Query

2

2 Searching a Substance by a Structure Query

What you will learn:

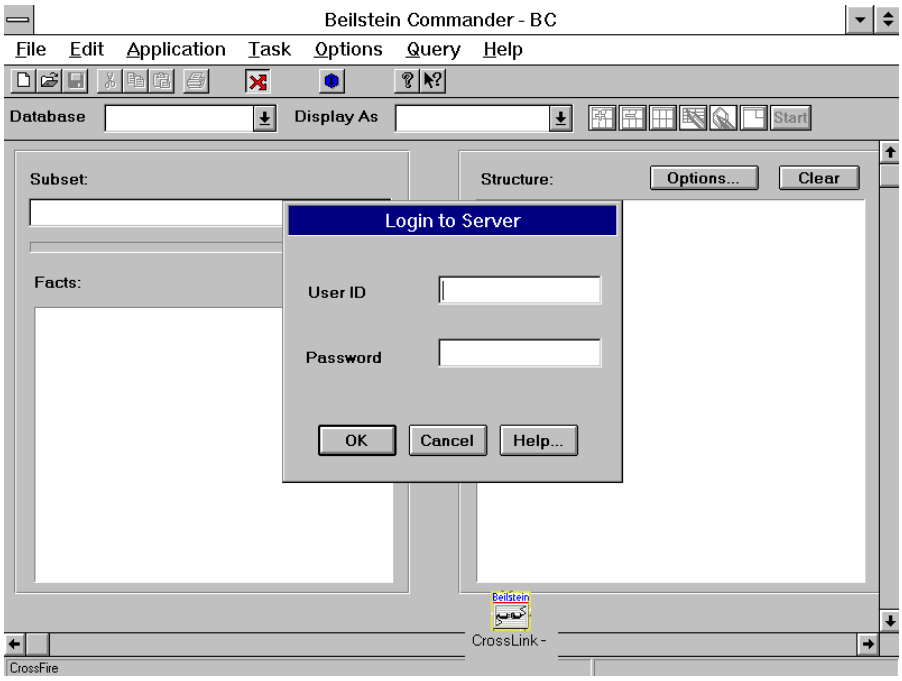
- ***Opening the Commander***
- ***Logging onto the server***
- ***Drawing full structures***
- ***Performing a search***
- ***Viewing the results***

2.1 Getting Started

Open the Commander which controls access to the whole CrossFire system by double click on the Commander Icon:

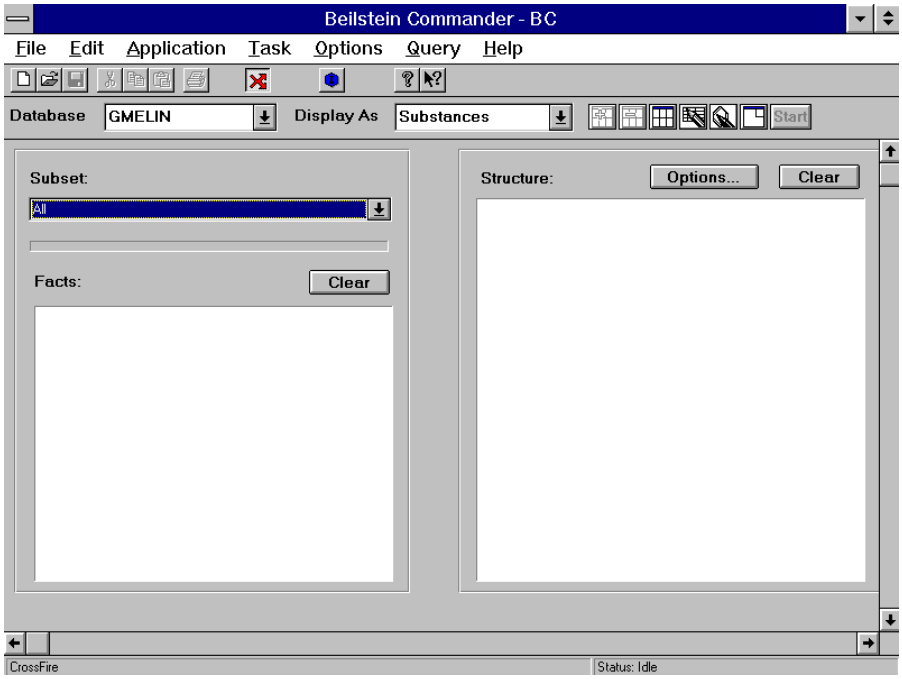
Activate the CrossFire system by a click on the CrossFire Icon  .

Fill in the appearing Login box with your User Name and Password:



Now the Commander is ready for you to perform searches and other actions.

Let us first have a closer look at the Commander:



On top you find a windows-type menu and below some shortcut buttons, including the CrossFire Icon you already used.

Keep your attention to the two list-boxes further down: the first allows you to select a database; within this tutorial we will focus on the Gmelin Database; make sure to select it now. The second box ("Display As") allows you to define the context for your searches, which will be dealt with later, we do not need to bother about it now.

Within this bar you also find more shortcut buttons, of which the most important is placed on the right: Start It is always used to start a search, e. g. to submit your query to the powerful search engine at the server. With the other buttons you can switch to other tasks within the Crossfire System (Structure and Fact Editor, Display Hits).

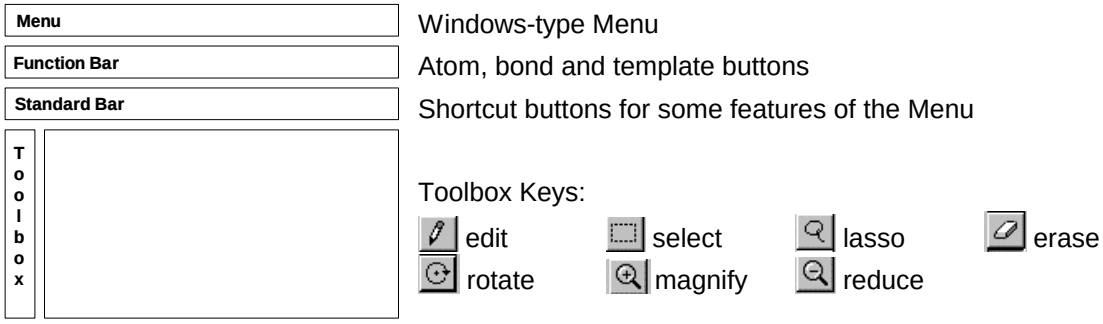
Starting a search is only possible from the Commander; from all other tasks you can switch to the Commander by the button (Beilstein Commander) or the function key F7.

The two empty windows within the Commander refer to the two important methods of formulating your query: the Structure and Fact Editor. On the right hand you see the Structure Editor Window, which we are going to open now by one of the following ways:

- click the Structure Editor Icon (benzene and pencil) on the button bar or
- choose "Structure Editor" from the Task menu or
- simply double click the window

2.2 Drawing Molecules in the Structure Editor

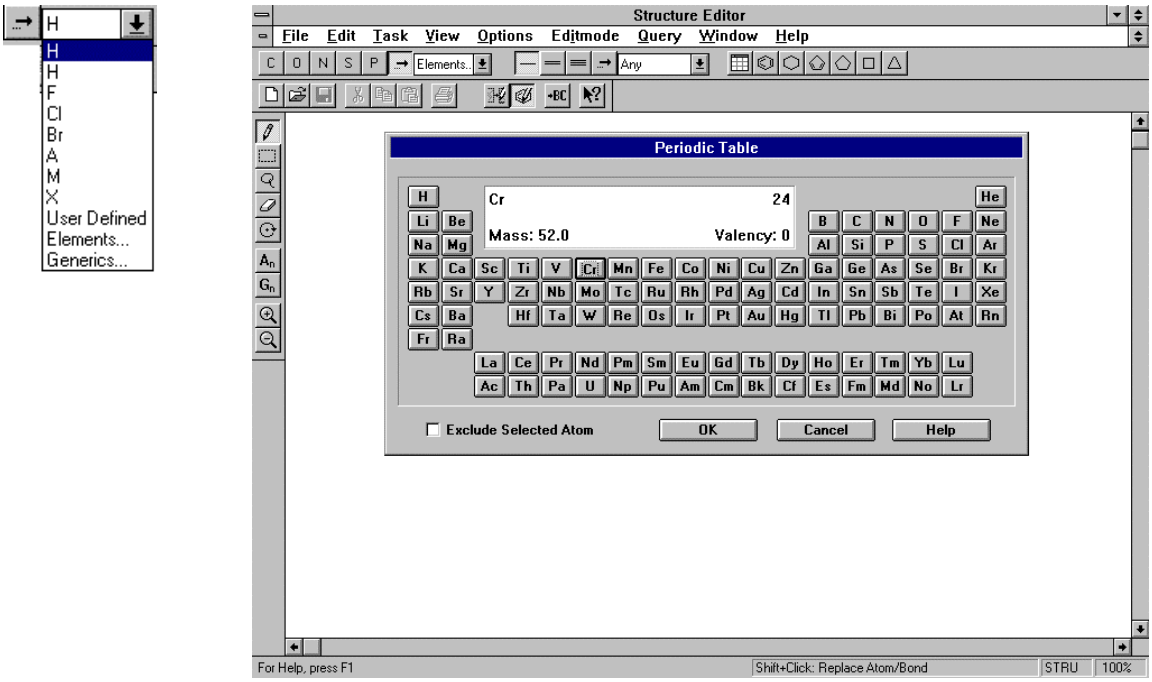
The following symbolic representation of the structure editor shows the layout in general terms:



In the first example we are going to search for the simple inorganic salt lead dichromate $PbCr_2O_7$ to get used to the way of drawing structures.

After opening the structure editor you will see that the mouse cursor has turned into a pencil (waiting for you to draw); this is called the edit mode. It is automatically selected once you open the Structure Editor. If you want to switch back to this edit mode later (after you have performed actions in another mode) click the pencil button within the toolbox

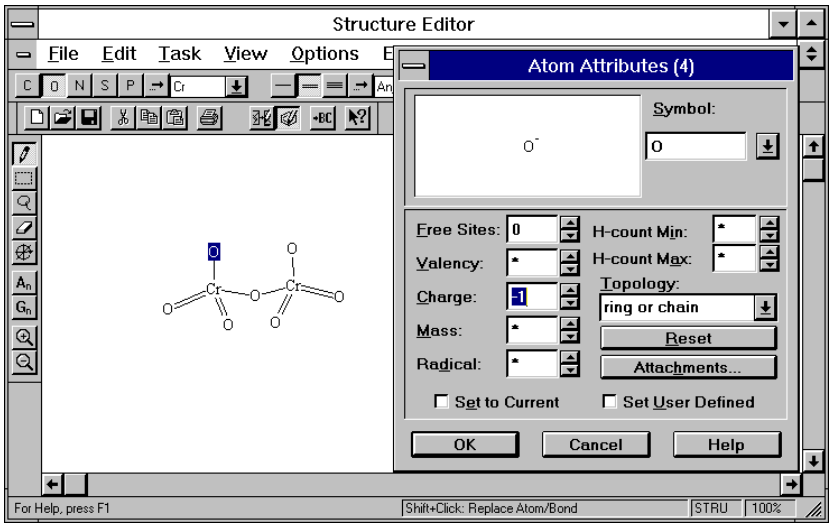
Now let us start to draw lead dichromate $PbCr_2O_7$. We are going to draw a chromium atom first. Choose the arrow box which is situated to the right of the atom buttons, select "Elements..." and pick Cr from the periodic table:



You will notice that the atom symbol in the list box has changed to Cr; it will be active until you choose some other element.

Click at the position within the Structure Editor window where you want to place the Cr-atom and it will be pasted. Now change the atom tool to Oxygen and the bond tool to double bond. When you now move the cursor (pencil) near the Cr-atom, you will notice that a small A appears in the middle of the pencil which indicates that you can connect a bond to this atom now. Click the mouse and move it, the double bond becomes visible and when you release the mouse an O-atom appears at the end of the bond. Repeat this action for a further O-atom. The other two O-atoms have to be connected by a single bond; select the single bond tool and draw the atoms. Now click again the arrow to the left of the atom list box to select the Cr-atom, connect it to the oxygen, change the atom tool again to O and draw the rest of the dichromate-ion in the way described previously. Now you

have drawn the atom skeleton, but you still have to set the charges to the O-atoms. For this purpose click at the atom where you want to set the charge; make sure to have selected the edit mode ( ). An atom dialog box appears, where you can set several atom attributes:



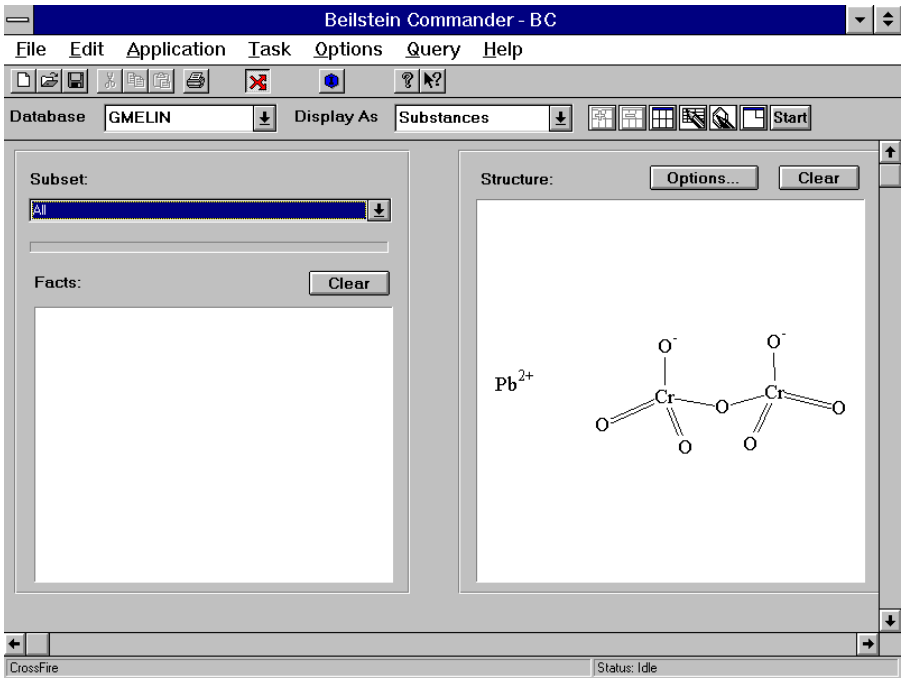
Set the charge to -1 and click OK; repeat this action for the other O-atom. Now the dichromate-anion is finished.

To draw the lead atom use again the arrow right to the atom list box, choose elements and pick Pb from the periodic table. Place the Pb somewhere in the structure window. Set the charge by pointing at the Pb-atom (edit mode) and choosing Charge +2 in the Atom Attributes box.

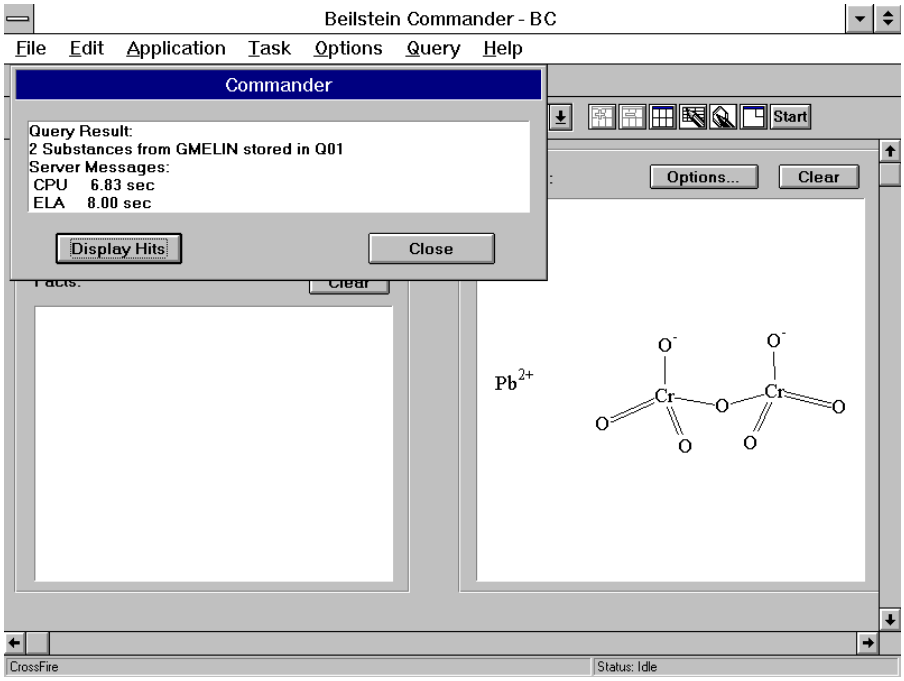
Note: Do not draw a bond between the lead and the dichromate-anion, salts are usually drawn in ionic form, i. e. as separate fragments with located charges, and never with covalent bonds

2.3 Performing a Search

After completing the structure you can transfer it to the Commander by using the F7-key or the BC-button. The Structure Editor window in the Commander should now look something like:

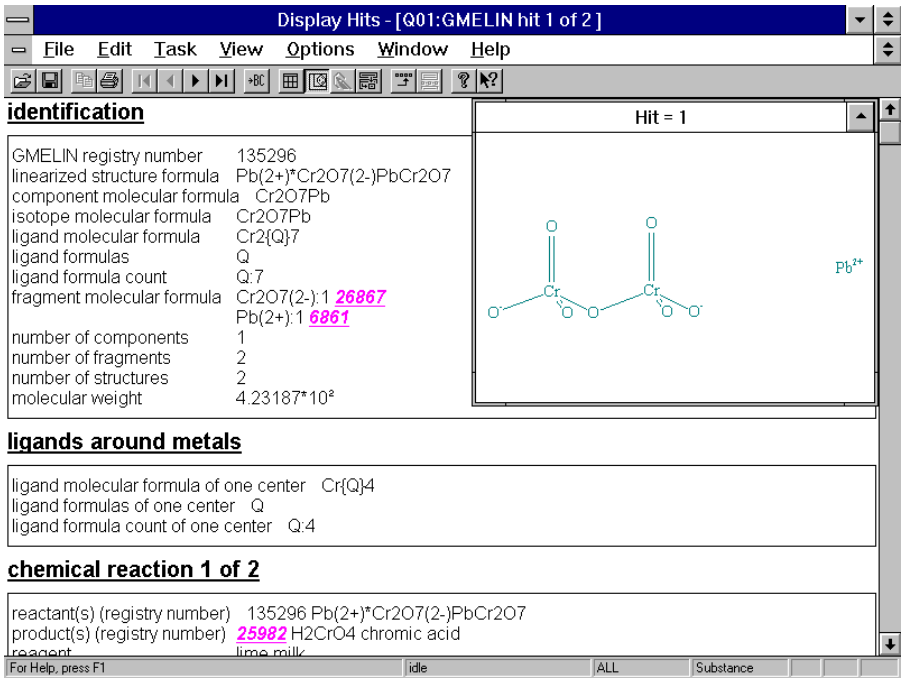


Press the "Start"-button and after a short while you get the answer of the system, which looks like:



2.4 Viewing the Results

You see that you obtained two hits in your search. Clicking Display Hits or hitting the Enter-key opens the Display Hits window with the first of the two hits:



You can browse through the record with the scroll bar on the right, additionally you can see a list of all available fields of this record by:

- Clicking the button 

  or

- Selecting "Field Availability" from the View menu

You can browse that list and select a field by double click for immediate positioning in the record.

You can switch the structure on and off by using the  button. If you are using a Macintosh as client, this is the way to change between the record and the structure of a substance (as mentioned in the introduction the Macintosh system does not allow to have a "Child-window" in the display).

Now let us have a look at the second hit. To switch from one hit to another you can use either the Edit menu or the arrow buttons:

- Choose the menu "Edit" and select "Go to hit ": a submenu allows you to select the first, previous, next or last hit of the current hitset; you can also jump to any hit of the hitset by choosing "Select hit number" and typing in the number of the hit you want to switch to

- Use the buttons  to move through the hitset; click the button  to select a certain hitnumber

When you now switch to the second hit in one of the ways explained above, you will see that this is the hydrate of $PbCr_2O_7$, which is also retrieved in this search.

This can be understood by the default settings of the structure query options, which allow for multicomponent compounds, radicals and charged compounds to be retrieved (for details about defining structure query options see Chapter 8).

At the beginning of each record you will notice quite a lot of fields for the identification of the substance, which become very important in the records of coordination compounds. They are also very useful to define a query for those compounds as they allow a classification of ligands and the kind of their bonding. We will perform such a search later, but we should now have a look at the most important fields of the identification data:

Component Molecular Formula (MF): gives a molecular formula in the Hill order; if the compound consists of more than one component (see below), a separate molecular formula for each component is given.

Isotope Molecular Formula (IMF): gives a molecular formula as above, but with additional information about number and kind of isotopes contained in the substance.

The fields **Ligand Molecular Formula**, **Ligand Formulas** and **Ligand Formula Count** detail the composition of a coordination compound and the kind of ligands (see chapter 4).

Fragment Molecular Formula (FRAGMF): gives a molecular formula for each fragment of the substance, e.g. two separate ions (including their charge) as in this example of lead dichromate.

Number of Components (NC), **Number of Fragments (NFRAG)** and **Number of Structures (NSTRUC)** give the corresponding numbers.

Note: not all the substances in Gmelin have a structure (i.e. connection table represented as drawing), for example salts such as NaCl or alloys do not have a structure.

Chemical Name (CN): The names given do not necessarily conform to the IUPAC-nomenclature. For a given substance there can be more than one chemical name, which also can be assigned in different languages, but not all compounds in the database have a chemical name.

Note: you can find a description of all fields - not only of the identification section (including the type of indexing) - when you choose within the Fact Editor from the "Help"-menu: Gmelin Reference Guide

By performing this first search and viewing the results we have got to know the general outline how the CrossFire system works and with every following chapter we will learn different ways of extracting information out of the database and by this being able to retrieve and view exactly the information we seek.

Searching a Substance by a Fact Query

3

3 Searching a Substance by a Fact Query

What you will learn:

- ***Using the Fact Editor***
- ***Employing the expand function***
- ***Using Hyperlinks to get additional information related to the actual hit***

3.1 Formulating a Factual Query

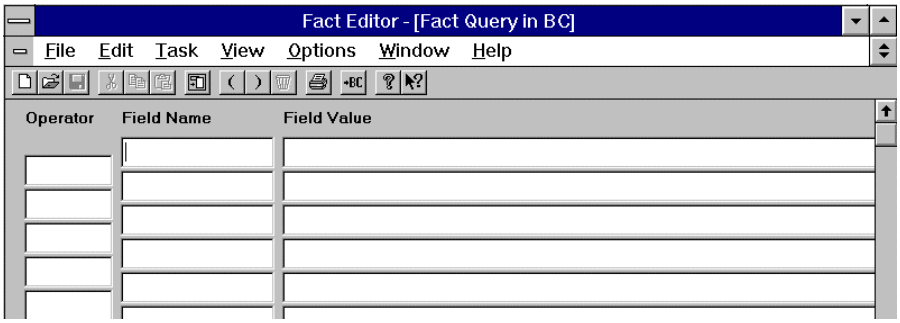
If you want to look for a substance there are several other ways than doing a structure query. You can search for all the fields within the identification section; sometimes it is useful to search for a molecular formula (as detailed in the previous chapter). There are special ways to look for organometallic compounds by the ligand descriptions, and you can also search chemical names (keep in mind that not all substances have one).

Before we can perform our next search we have to clear the screen of the Commander by clicking the Clear button of the Structure Editor window.

In this example we will search $K_4[Fe(CN)_6]$ by the component molecular formula.

The substance consists of one component and two different fragments, i.e. four times the potassium-cation and one hexacyanoferrate-anion. To formulate the query, we will open the Fact Editor by one of the following ways:

- click the Fact Editor Icon (table and pencil ) on the button bar or
- choose "Fact Editor" from the Task menu or
- simply double click the window



You will notice that the Fact Editor consists of three columns:

Field Name: to fill in the code of the field (i.e. property) you want to search

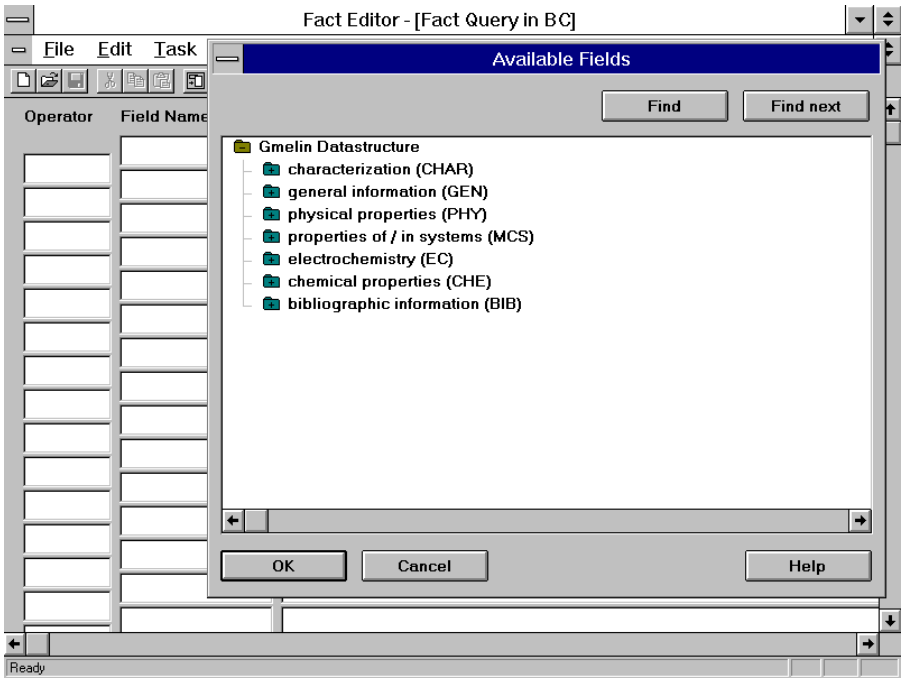
Field Value: to fill in the term(s) (e.g. values, keywords, phrases) you are looking for

Operator: to combine two or more queries by logical operators (and, or, not, proximity)

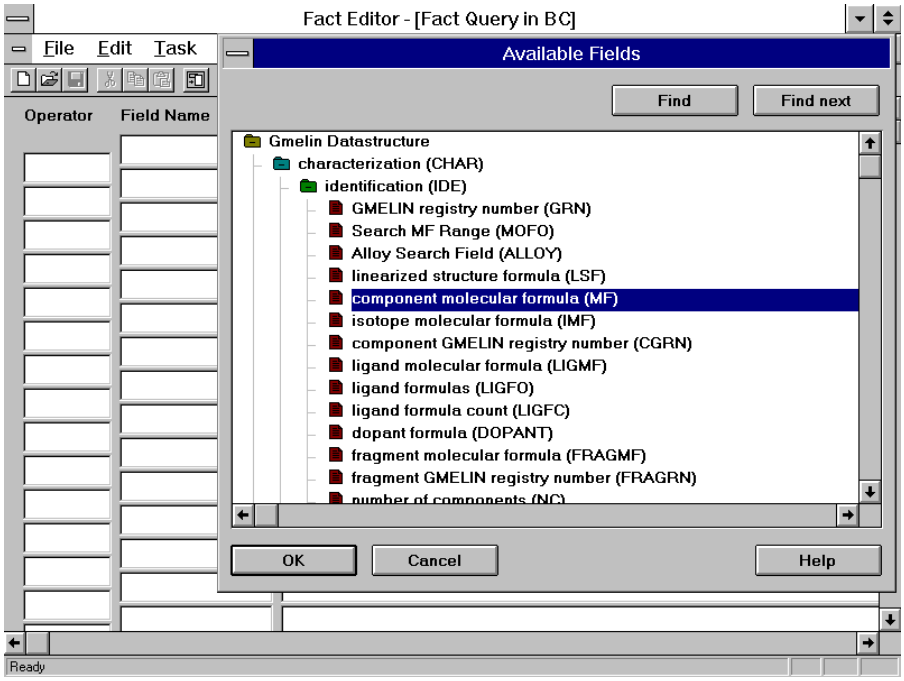
After opening the Fact Editor the cursor is automatically placed in the first Field Name cell of the table, ready for you to fill in the code. There are three ways of doing this:

- Pick the desired codes from the data structure, which can be opened by the -button (List Values) or the F2-key. Select the appropriate field code by double click.
- Within the "Available Fields" dialog box (see screenshot below) you can search for a special field name by using the "Find"-button and typing the name of the desired property or attribute. This is very useful if you do not know in which section of the datastructure a given property is situated.
- Type the codes directly into the fields (if you already know them)

After clicking the ("List values"-button) the "Available Fields" dialog box appears, where you can see the seven main sections of the Gmelin Datastructure.

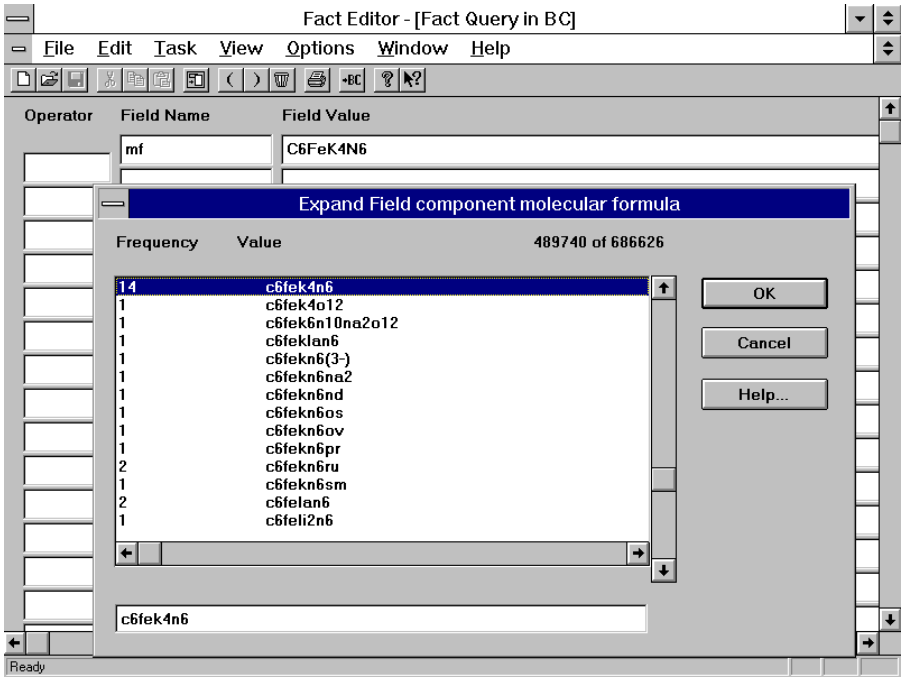


You can view the sub-sections and fields of the main section by opening the folders. The field "Component Molecular Formula" is to be found in the section "Characterization" and the sub-section "Identification" as you can see from the next screenshot:



As mentioned above, if you do not know in which section a certain property is listed, you can always use the "Find"-button and type in the name of the property or some relevant characters of it. The highlight will then automatically be placed at the desired entry of the datastructure.

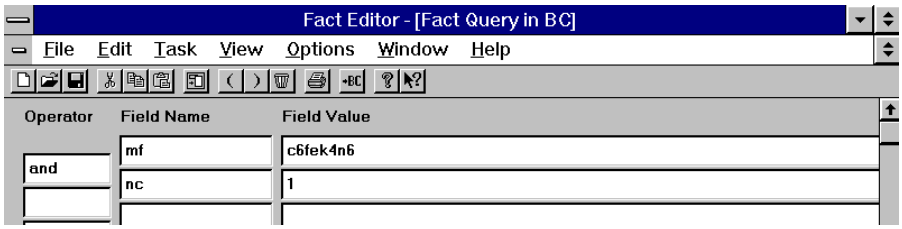
Click OK and the corresponding field code (MF) is placed in the Field Name cell. Now use the Tab-key to switch to the next column to fill in the value, in this case the molecular formula. Remember to use the Hill order[1], i.e. C6FeK4N6. You can use the "Expand"-function (List Values) to see an index list of all data and their occurrences in this field. You can browse the list and pick the entries. If you look at the index list now, you might wonder about the high number of occurrences of this particular molecular formula:



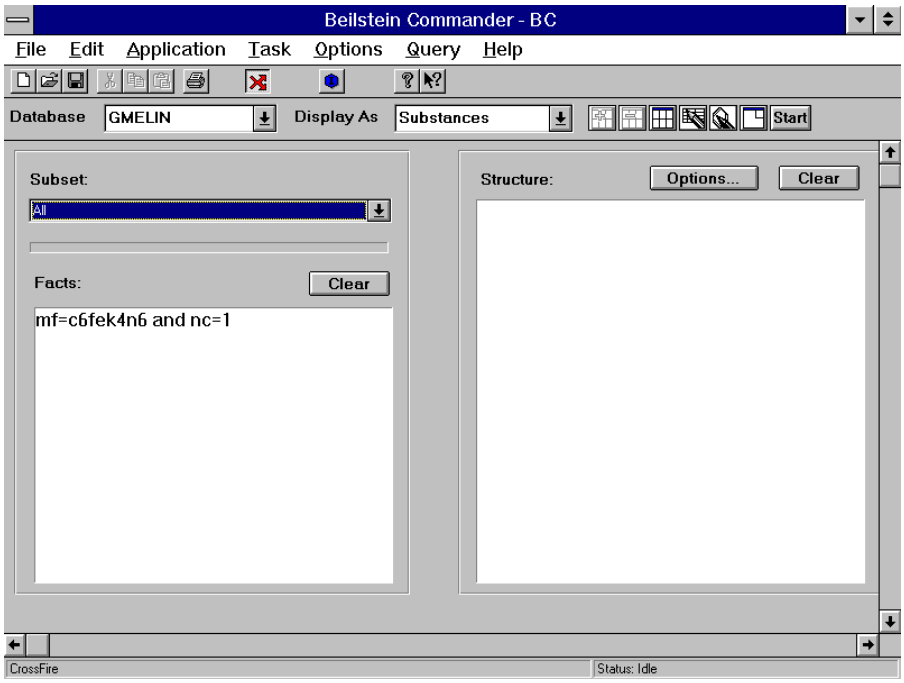
Keep in mind that in this field all isotopes of an atom are combined in one formula. Furthermore you will expect to find some multi-component compounds with a $K_4[Fe(CN)_6]$-unit. As we are interested in the single-component compound we will specify this fact in the query: Use the Tab-key to place the cursor in the first operator cell. We want to retrieve substances which fulfil both of our criteria, therefore the queries have to be combined with the operator "and", which we are going to type in now. You could also use the "Expand"-function to view the list of all available operators and pick "and" from that list. Move the cursor to the next

[1]The molecular formula is listed in the following systematic way: First C, then H followed by all other atom symbols in alphabetical order

"Field Name"-cell, open the "Available Fields" dialog box by the -button or the F2-key and select the code for "Number of Components", which can be found in the identification section, too. Alternatively, you could type "NC" directly into the cell. Now move the cursor to the "Field Value" cell and type in "1". The following screenshot shows how the whole query looks like:

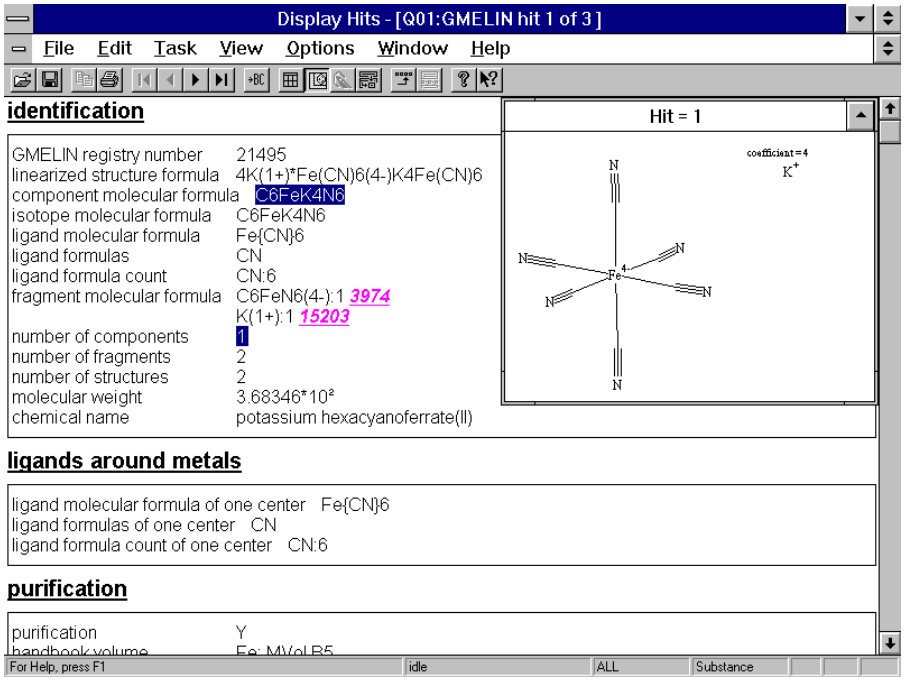


Now switch back to the Commander by the BC-button or the F7-key. It should look like:



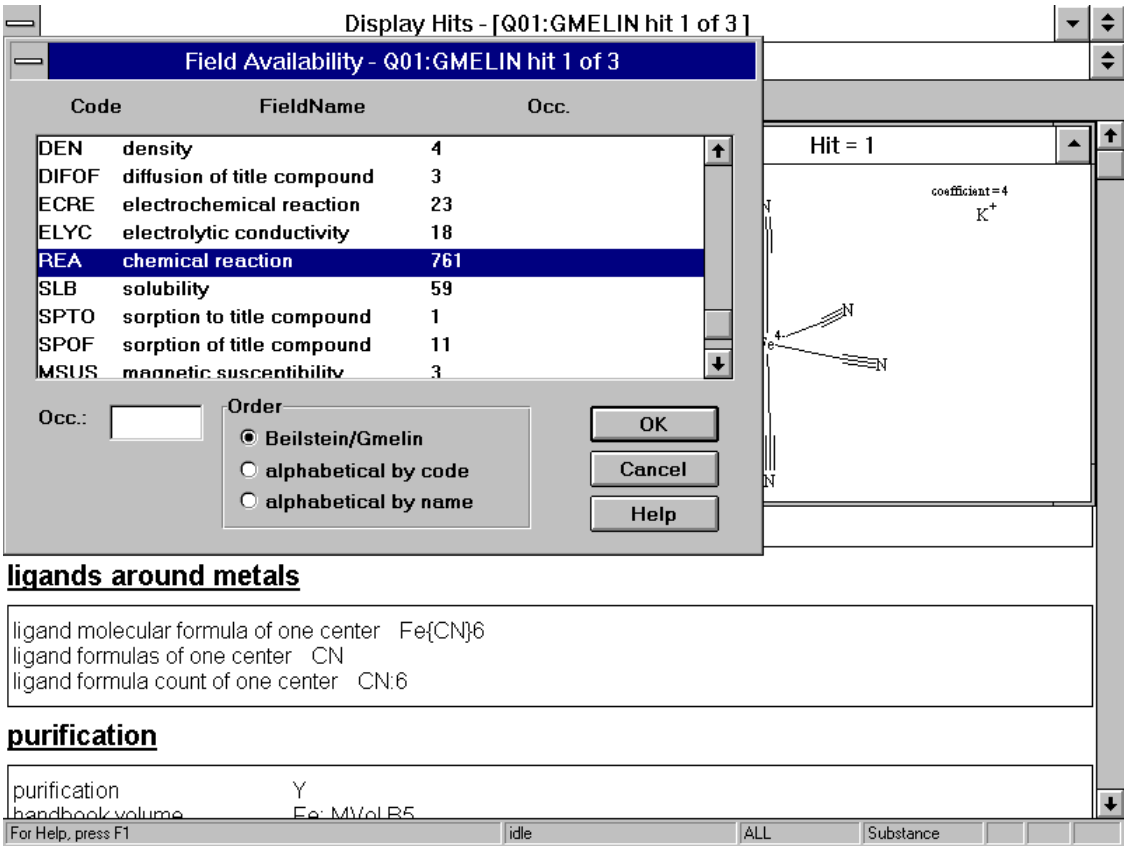
Start the search with the "Start"-button or the F7-key. Again you will get the answer of the system in an additional window, where you can click "Display Hits" to see the first of the three hits.

3.2 Moving about the Display



In the display above you see the search terms "Component Molecular Formula" and "Number of Components" highlighted. If you do not have a highlight in your display go to the "View" menu and select "Highlight Hits". Alternatively, you can switch off the highlight in the same way, if you do not like it (If your display looks altogether differently from the screenshot shown here, select "All Fields" in the View menu).

You will also notice the coloured underlined numbers written after the fragment molecular formula. They indicate a hyperlink to the respective ions, which also have their own record, where you can get more information. Before trying the hyperlinks let us have a closer look at the whole record. Use the "Field Availability" command to see the list of properties (as described in the previous chapter).



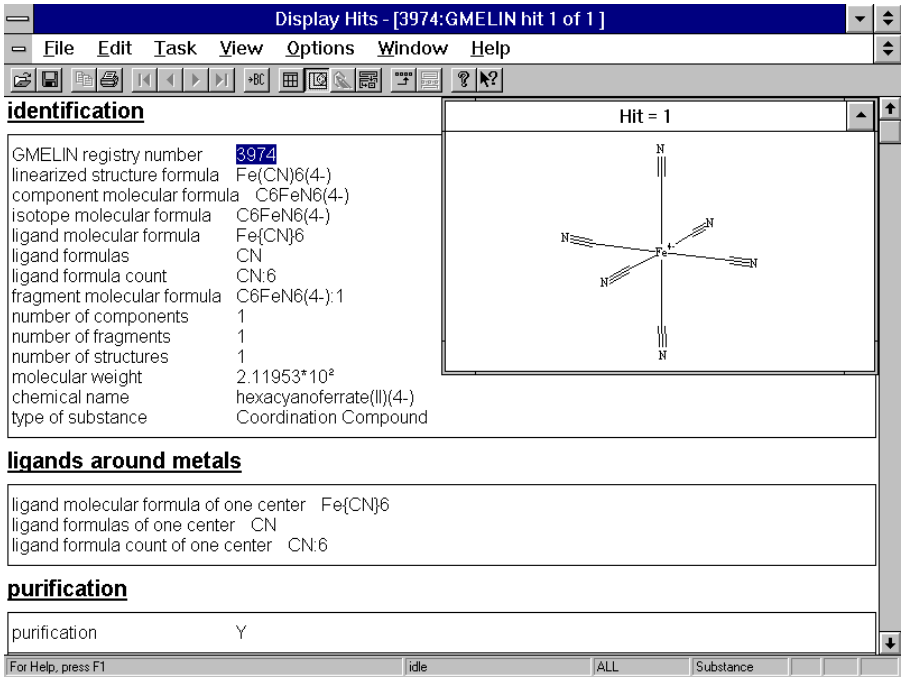
As you would expect, a lot of information is collected about such a well known compound. Especially in such long records it is very useful to employ the Field Availability list to move about the record by selecting the desired field by double click for immediate positioning in the record. Call up the Field Availability list again to jump to other data of the record. You can also define the order of the Field Availability list by the options at the bottom: the default setting is according to the hierarchical Gmelin datastructure, but you can also choose an alphabetical order of either field names or field codes. If you move to the field reactions you will notice 761 entries; this number is not surprising when you realise that all reactions in which $K_4[Fe(CN)_6]$ takes part are listed, not only preparations of $K_4[Fe(CN)_6]$, but also the many reactions in which $K_4[Fe(CN)_6]$ acts as a reactant. Later in this tutorial we will learn ways of specifying the query so as to retrieve only preparations or to exclude them.

Another feature of the FA-list will help you to deal with this high number of chemical reactions. On the bottom left of the FA-list you see an empty field named "Occ.:", where you can fill in to which occurrence of the selected property you want to jump.

After this bit of theory let us practice moving about the record.

3.3 Hyperlink to a Fragment

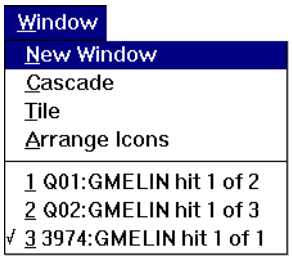
Click the coloured underlined Gmelin Registry Number *3974* of the hexacyano ferrate (IV) ion and a new window with the record of $[Fe(CN)_6]^{4-}$ is opened which is rather similar in outer appearance to the substance record of $K_4[Fe(CN)_6]$.



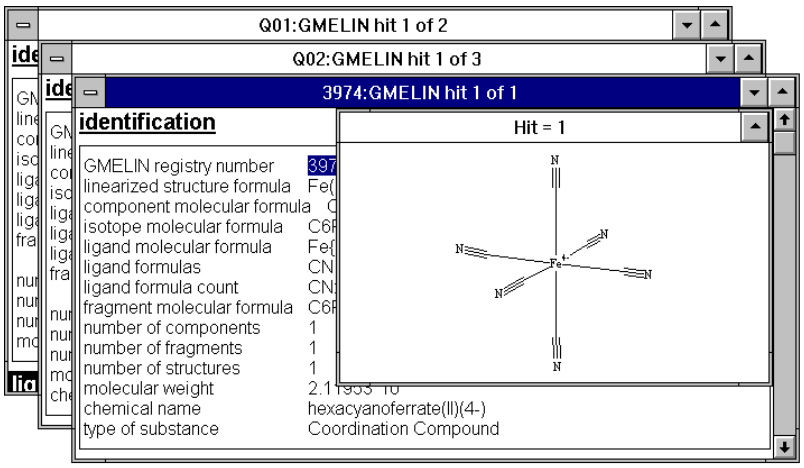
If you want to know more about the data of this ion move about the record using either the scroll-bar or the FA-list. After viewing the record we want to switch back the $K_4[Fe(CN)_6]$ record.

There are several ways to do this:

- Open the menu "Window"; at the bottom of the listbox you will find a list of all the records which are currently opened and the active one (in the front) checked. Click the hitset of $K_4[Fe(CN)_6]$ (you can recognise it by the number of hits which is 3) and the corresponding record will appear on the screen. The record of the hexacyano ferrate (IV) ion remains opened and is only shifted to the background, but can be viewed again any time during the session.



- Alternatively, you can select "Cascade" from the menu "Window" and you will see a cascade of the headers of all opened records. Then you can pick the record you want to switch to. With this procedure the hexacyano ferrate (IV) ion record remains opened, too.



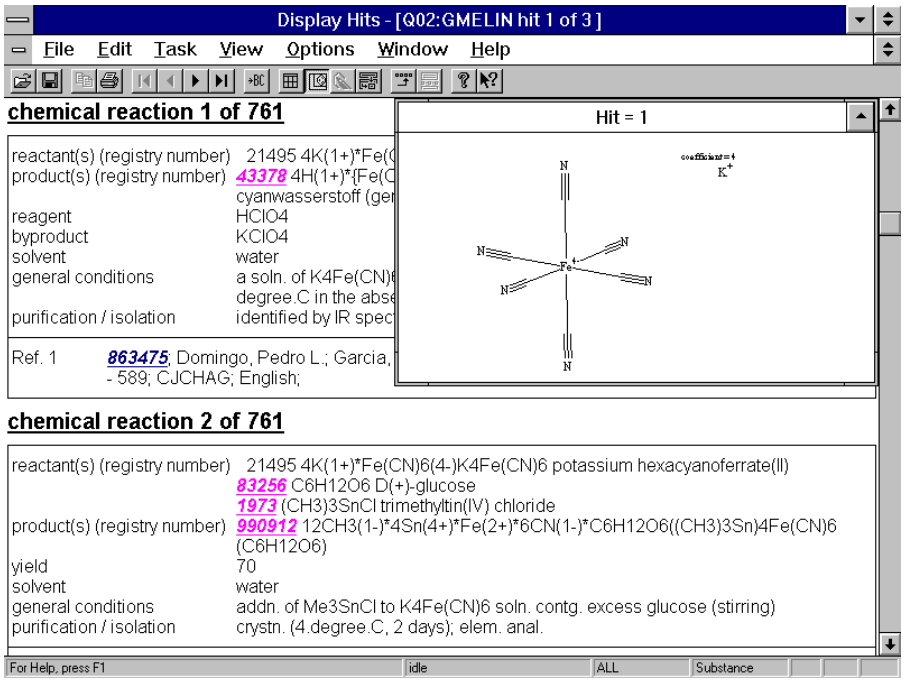
- Close the currently active window and you will be placed again in the record of $K_4[Fe(CN)_6]$. By this procedure the record of the hexacyano ferrate (IV) ion is closed but you can always look it up later during your session by using the hyperlink again.

It is recommended to close all the records which you probably will not need again for two reasons: first, after a while you will presumably get confused with a large number of windows and will not find the one you look for again; secondly, the more windows are opened at one time, the more memory is needed and the PC will be considerably slower.

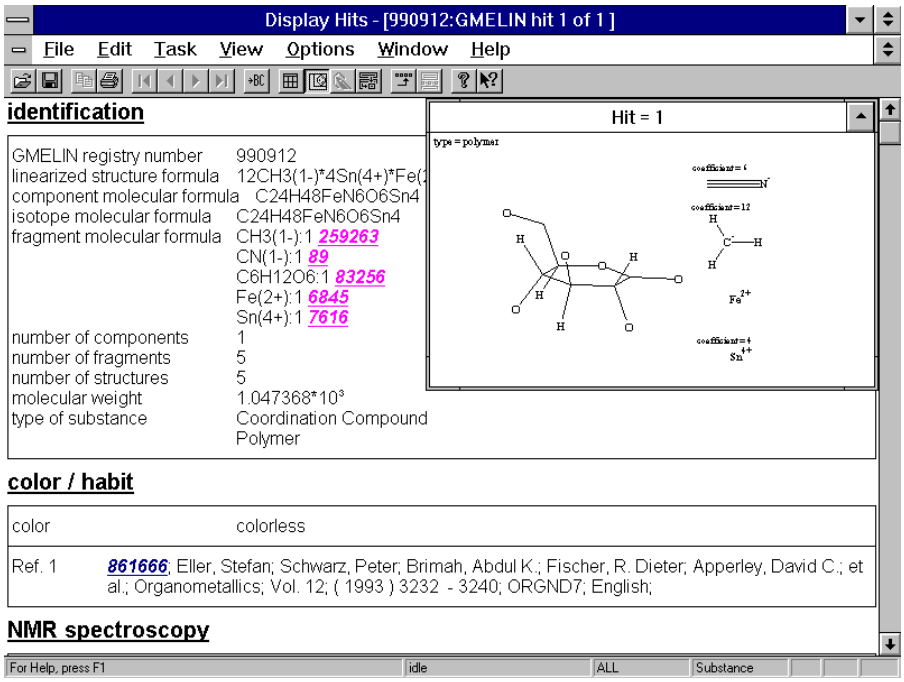
All the records retrieved as hitsets (which can be distinguished by numbers like Q01, Q02...) will be temporarily saved while you are logged in and can be called up any time during your session. Later in this tutorial we will re-open the hitset we just retrieved (in this tutorial it has the number Q02, your number might be different, depending on how many searches you have done before).

3.4 Reaction Records and Substance Hyperlinks

Let us view the first reactions of $K_4[Fe(CN)_6]$. Just double click the chemical reactions entry in the FA-list. You will be placed now at the first entry.



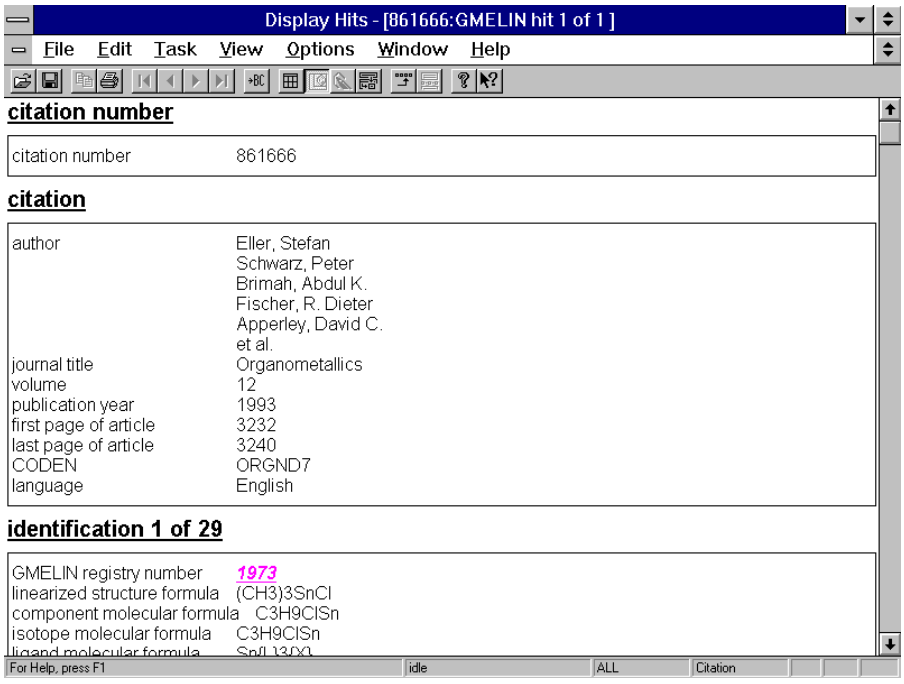
If you look closer at the second reaction entry you will observe hyperlinks to two reactants and the product of the reaction. You might find this entry rather interesting as it leads to a multicomponent organometallic compound, but from the molecular formula you can hardly guess the structure. Click the hyperlink to the registry number of the product (*990912*) and you will see the corresponding record.



If you move the cursor further down to the reaction entry you will observe the same information as in the record of $K_4[Fe(CN)_6]$, because the same reaction data are stored in the records of both reactant(s) and product(s).

3.5 Hyperlink to a Citation

There are also hyperlinks to a different context, which we have not tried yet. You will notice that each reference has a citation registry number which provides a hyperlink to more information about the whole article. Let us now try the hyperlink to the reference of this reaction:



Thereby you can view the detailed bibliographic information as well as a list of all substances reported in the article. This is often a very effective way when you want to know if similar substances have been reported by the same author. Additionally, it allows you a totally different view of the data and you can use further hyperlinks to the substances mentioned.

By this procedure you might get ideas you would not have thought of before and you will regard this system of hyperlinks as a powerful creativity tool.

After experiencing all the different possibilities of hyperlinks let us not forget that we retrieved two more hits from our original query. We will have a quick look at them, too.

The second hit looks much the same as the first, but a different type of substance has been assigned for it by the author, so it has been given a different registry number, thus resulting in a different record. The third hit is an isotope labelled hexacyano ferrate containing ^{15}N atoms.

In the first two chapters you have learned to search for specific compounds by either structure or identification data. We will now move on to a search for groups of substances by defining substructure queries.

Substructure Search
Using the Structure Editor

4

4 Substructure Search Using the Structure Editor

What you will learn:

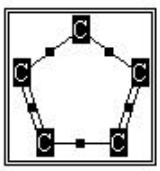
- ***Drawing coordination compounds***
- ***Using pre-defined generic groups and free sites***
- ***Using the different options of the display (full and short display , user-defined view)***
- ***Displaying hitsets in a different context***

4.1 Drawing a Coordination Compound

After now having learned about the most important ways of looking for a special substance and having experienced some fundamental features of the Display Hits we will now perform a substructure search. First of all we have to clear the Fact Editor window within the Commander, then we will open the Structure Editor.

We are interested in bis-alkynyl substituted titanocenes and we want to allow for any number and kind of substituents at the cyclopentadiene rings.

First we will draw the Cp_2Ti unit. Click the ring template  (cyclopentadiene) and the ring will be pasted at a random position within the Structure Editor window.

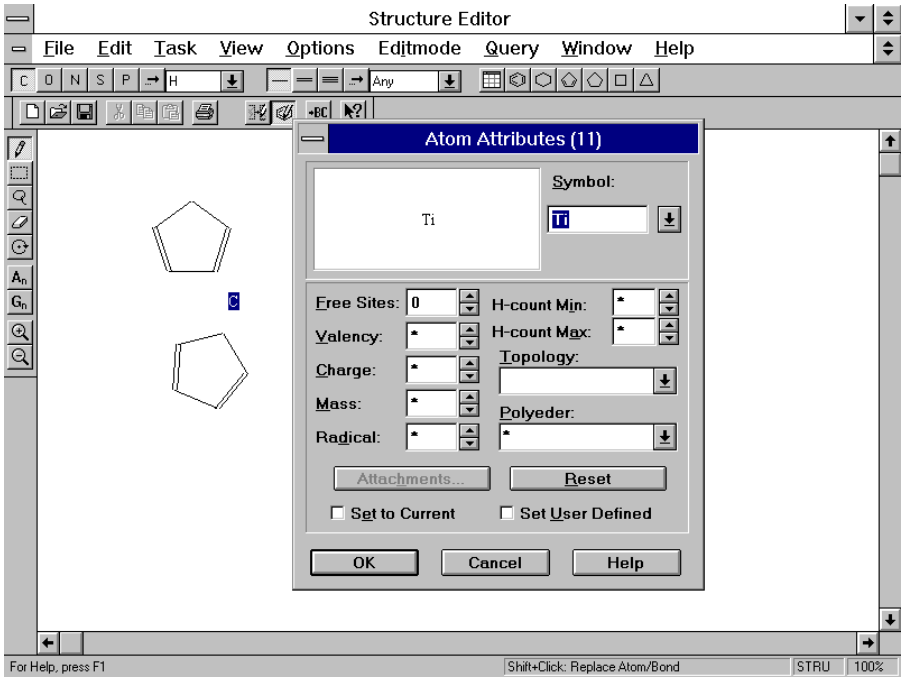


You will notice that now the ring is selected (all rings and bonds are marked with black squares and the whole ring is contained in a double-lined rectangular box). This means that a number of operations can be carried out on the selected object: you can grab it somewhere in the box and move it to any position. You can also carry out annelations and spiro-formations: for annelations you have to grab the square marking the bond and drag it to that bond of another fragment you want to merge it to; for drawing spiro-compounds you have to drag it by an atom.

Since we do not want to draw any annelated or spiro-compound now just move the Cp-ring to some adequate position. To draw the second Cp-ring you have two

possibilities: first, you can click the  -button again and the second ring will appear on the screen; secondly, you can copy the already drawn ring as long as it is selected by holding down the Shift key, grabbing the ring somewhere and moving the mouse to the position you want to place the second ring. The Shift+Drag method is especially useful for copying a more complex structure fragment.

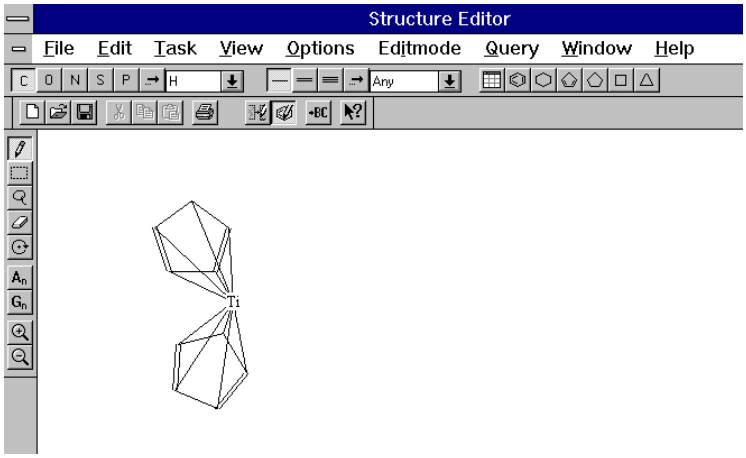
Now we will draw the Ti-atom: it can be done as described in the first chapter by choosing "Elements..." from the atom listbox and picking Ti from the periodic table. However, usually it is faster to draw any currently active atom in the place you would like to locate the Ti, then click that atom again and type the element symbol directly in the appearing "Atom Attributes" box followed by clicking OK or hitting the Enter key. This also permits you to define further atom specifications such as charges, radicals and free sites (see below) in one step.



The next question is: How do we draw the coordinative bonds between the metal and the ligands? Within the database all these bonds follow the same convention: always draw a single bond, in spite of exceeding the valency of any atom.

Sometimes this way of drawing might look strange and somehow "not right", but nevertheless it is an unambiguous representation which allows to take all interactions into account. It is also important to realize that there are never delocalized electrons in ligands like a Cp-ring and the alternating double and single bonds are explicitly drawn.

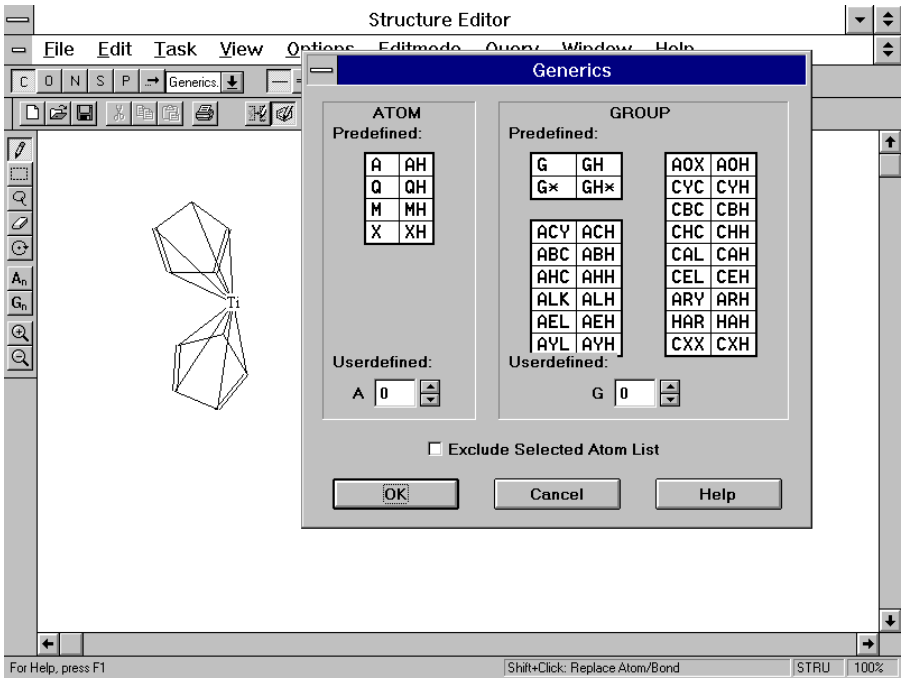
After drawing a bond between every carbon atom of the two Cp-rings and the Ti-atom the structure fragment should look like:



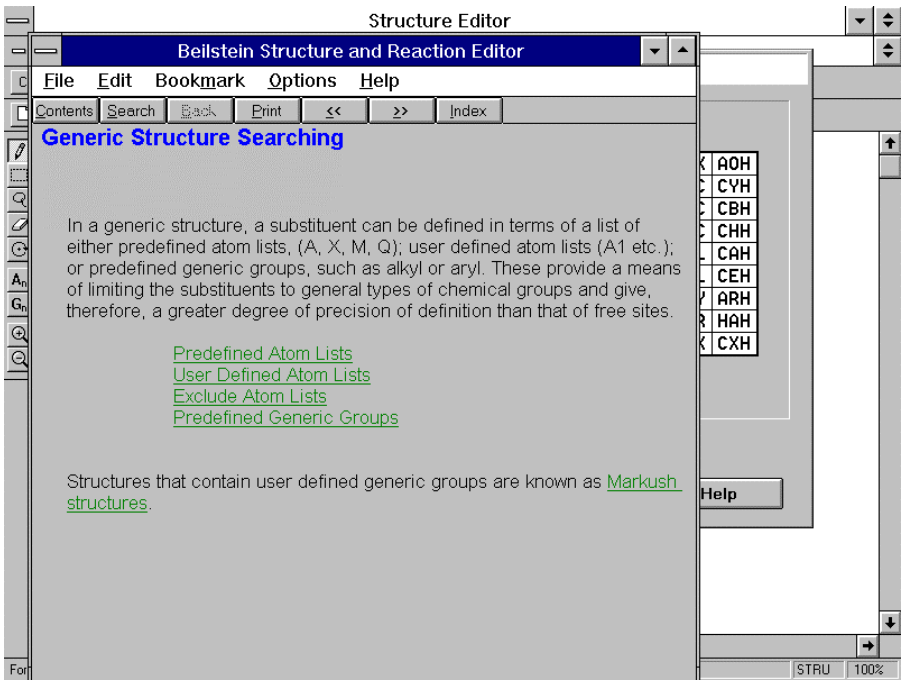
4.2 Using Pre-defined Generic Groups

The next step is to define an alkynyl substituent in a general form. The Structure Editor provides you with several pre-defined generic groups such as alkyl, aromatic, heterocyclic etc.

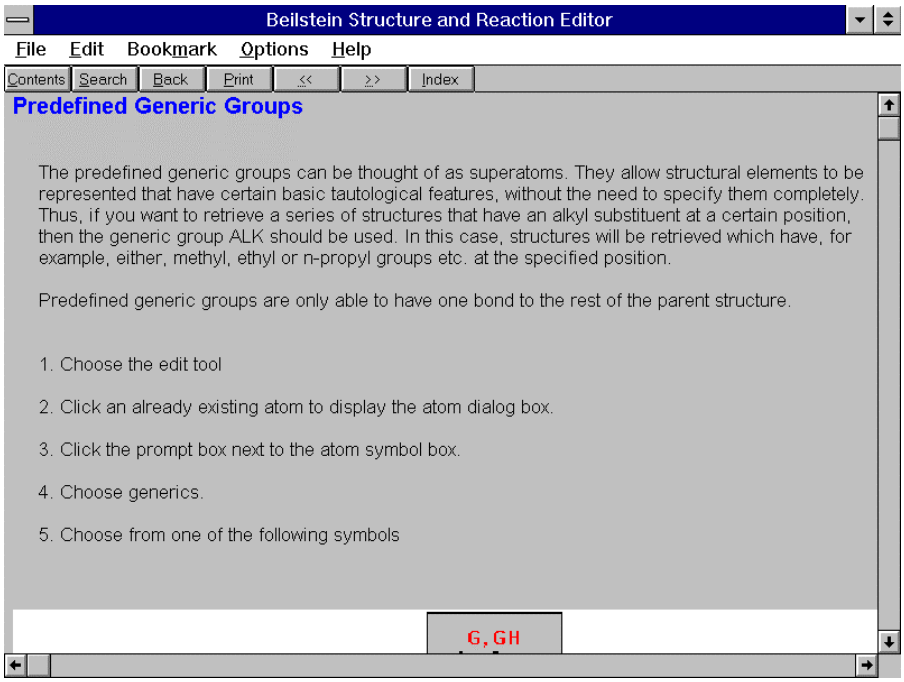
Choose from the atom list box "Generics..."; a dialog box with all kinds of Generics is opened; on the left you find pre-defined atom lists and a box to define user-defined atom lists. On the right hand are the generic groups: on top a table with pre-defined groups and at the bottom the possibility to select a number for a user-defined generic group.



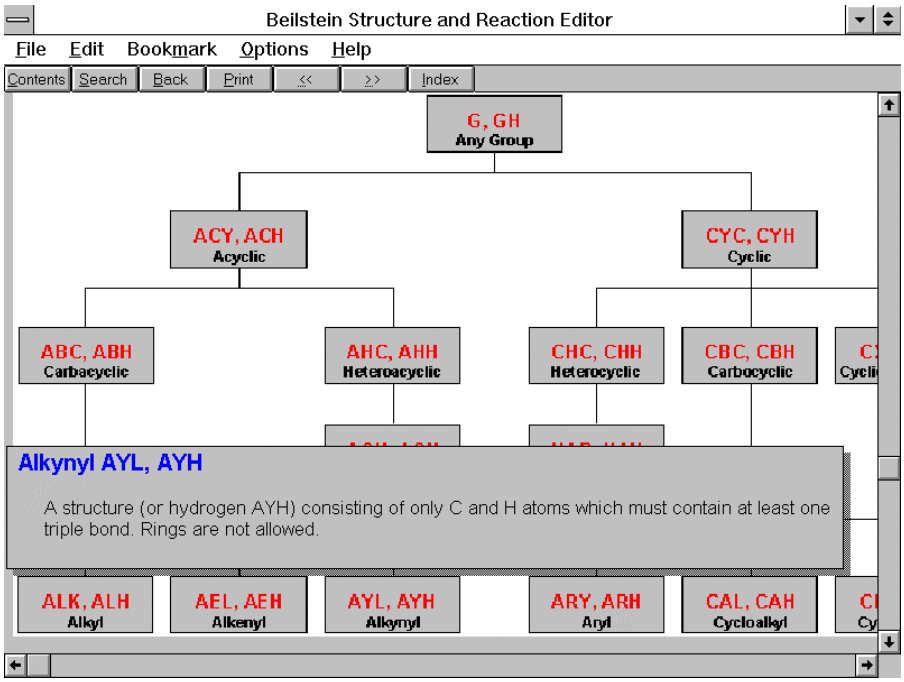
At first sight the abbreviations of the predefined generic groups look a bit confusing, but by employing the help-function there is no problem to use them. If you click the "Help" button at the bottom right, the Help-window with the contents of help on generics is opened:



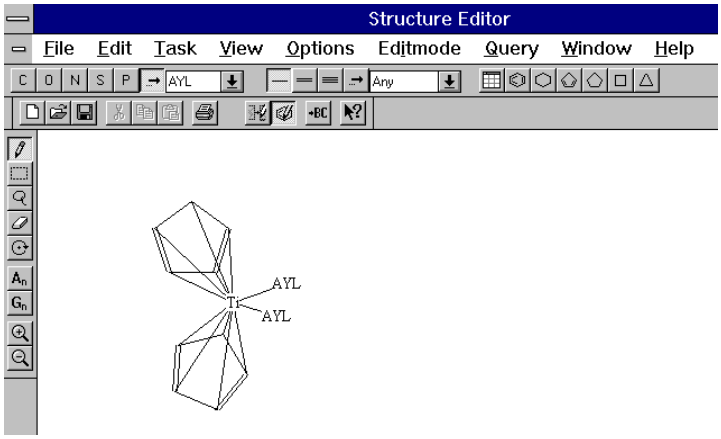
If you choose the last entry "Pre-defined Generic Groups" you will see the guide for using them. Maximize the help window for a better overview:



After a short introduction and a step-by-step instruction how to use them you find there a hierarchical tree of all pre-defined generic groups. Scroll down until you reach the bottom of the tree. You will notice that a every branch of the tree the group gets more specified, starting on top with the most general G = any substituent. You can open an additional window with the exact definition of each group by clicking at the respective codes of the groups. In case of the AYL-symbol it should look like:



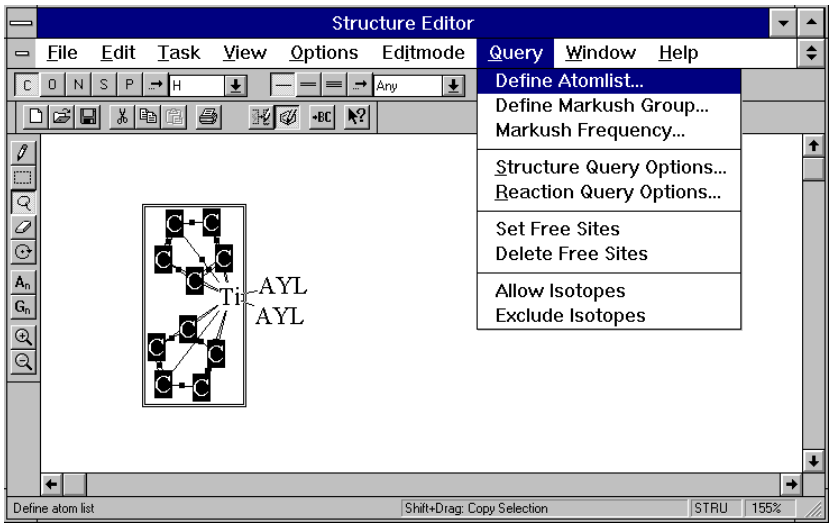
After having verified the definition of the rest you want to add to your drawing, close the help window and select the AYL-symbol in the small table of the pre-defined generic groups. This symbol is now placed in the list box next to the atom buttons and you can continue to draw the structure by adding two alkynyl-groups:

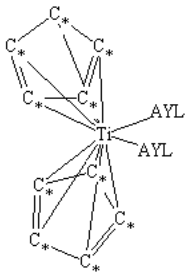


4.3 Setting Free Sites

The next step is to allow any kind and number of substituents at the Cp-rings, which is done by setting maximum free sites to all the carbon atoms. Of course

you can select every single atom one after another and set free sites for each atom separately, but within the "Options" menu you have the possibility to set maximum free sites to all the atoms you have previously selected. Entrap all atoms of the two Cp-rings with the lasso tool. and choose "Set Free Sites" in the "Options" menu. You will observe that all the C-atoms are now marked with asterisks to indicate the free sites setting.

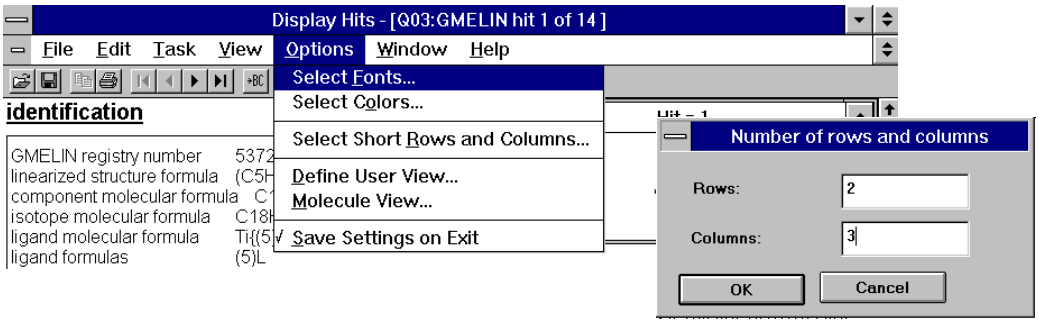




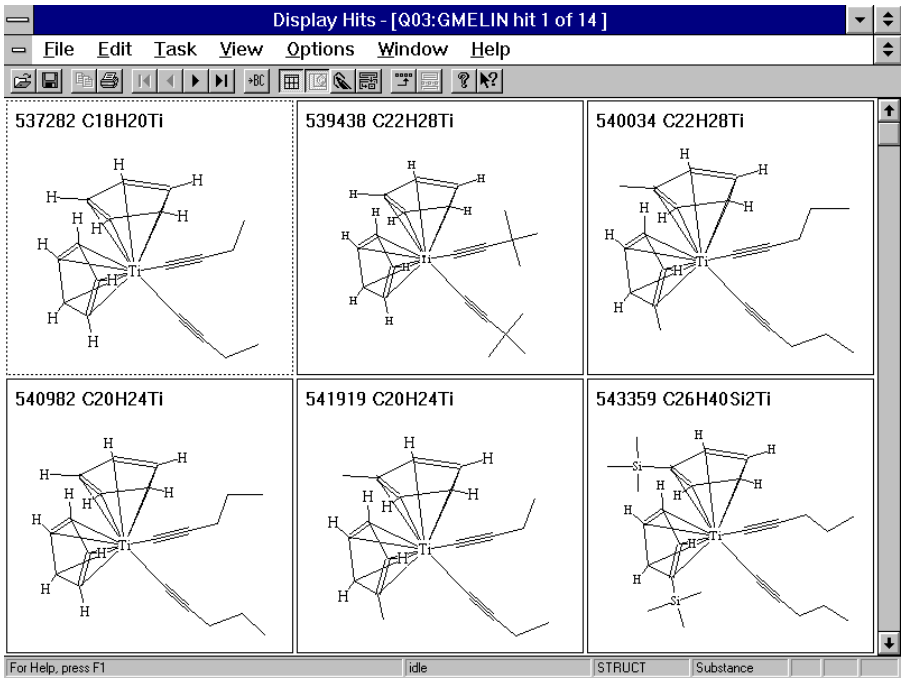
Now the query structure is complete and can be transferred to the Commander. Press the Start-button and you will get the result of 14 hits.

4.4 Using the Short Display

Switch to the Display Hits. For viewing a larger number of hits as in this case it is often useful to get an overview about the retrieved structures first before having a closer look at their data. This is best done by switching to the so-called Short Display, which can be selected from the "View" menu. You can define the number of structures displayed at a time by choosing "Short Rows and Columns" from the "Options" menu and typing in the number of rows and columns (up two four structures in each row and column can be chosen independently). Which is the best view depends on the size of your screen as well as the size of the molecules. We will now define two rows and three columns.



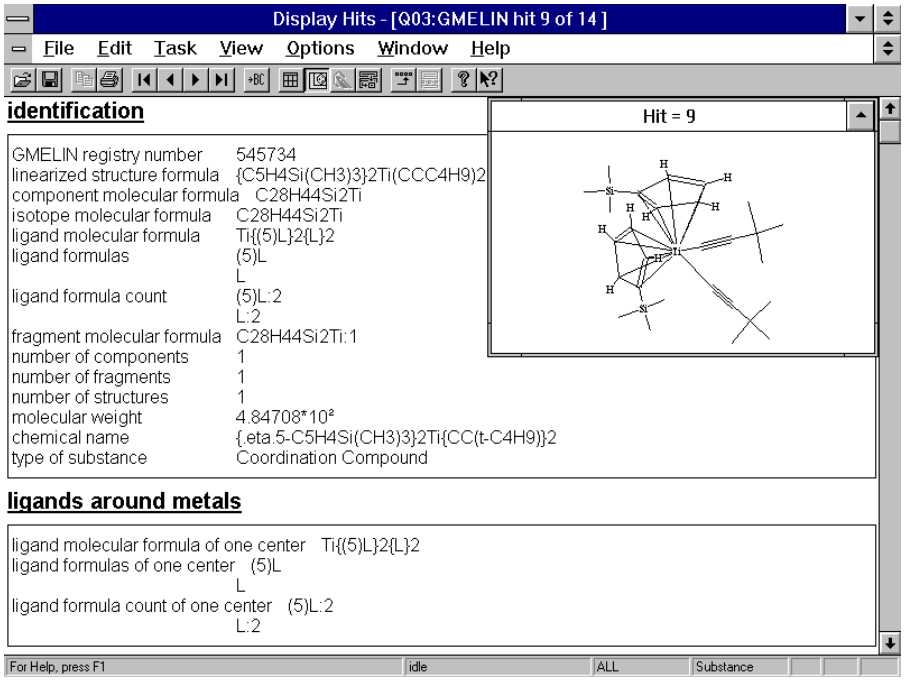
After you have defined the numbers the display does not automatically switch to the Short Display. You can now choose it either from the "View" menu or by the  button. This button also allows you to switch back to Full Display.



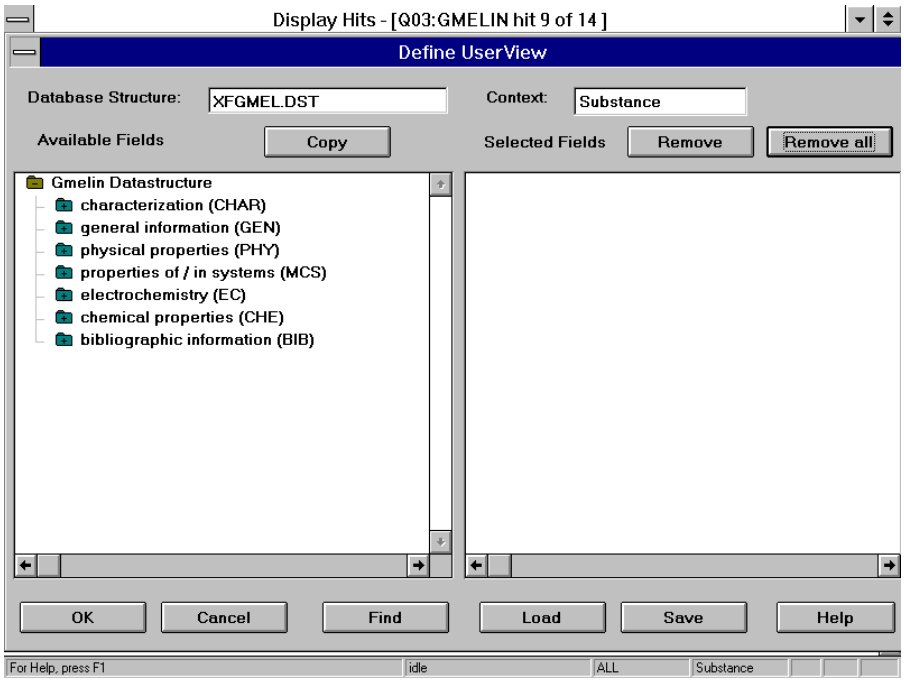
You can use either the scroll bar or the arrow keys to move about the short display. If you double click on a structure you will be placed in the full display of the corresponding compound. We will now take a closer look at the data of hit no. 9 in the full display and will learn about adjusting the display to our convenience by defining a "User-view". This allows you to have only those data displayed that you choose

4.5 Defining a User-view

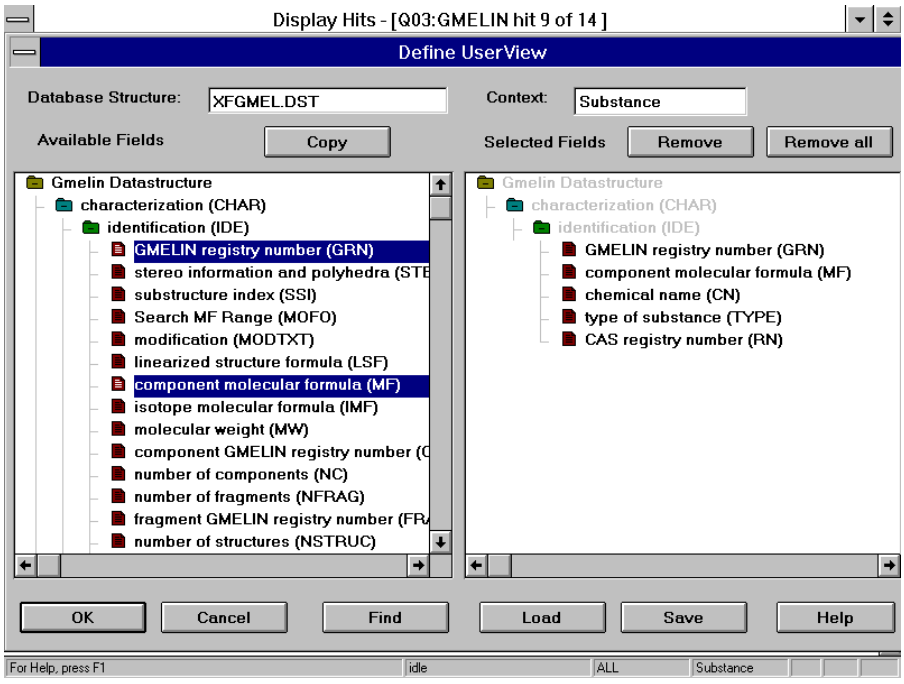
Double click hit 9



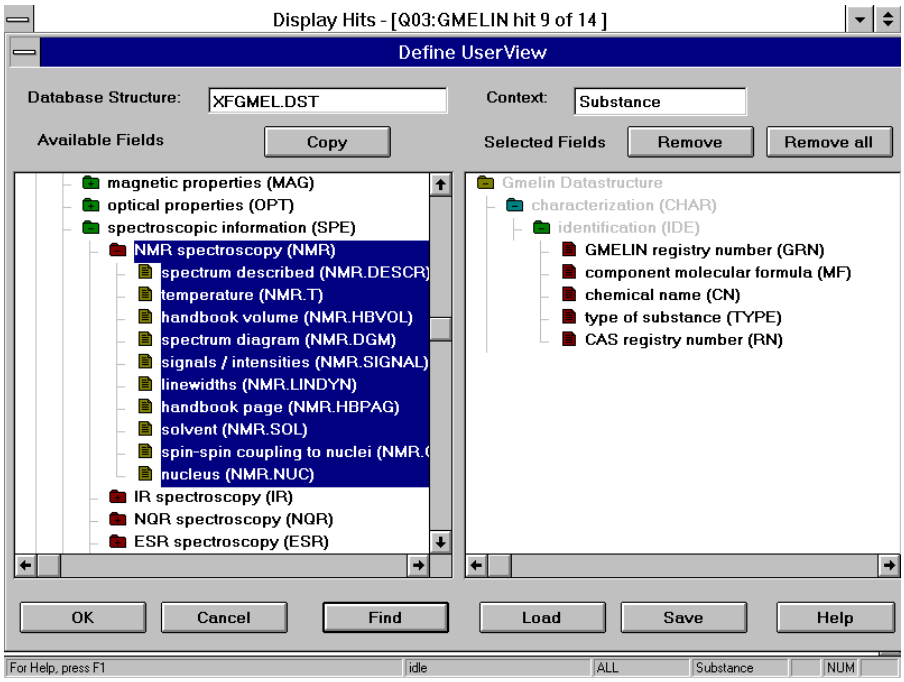
Click menu "Options" and select "Define User-View". You will see two separate windows, one on the left with the Gmelin datastructure (which looks the same as in the Fact Editor "Available Fields" box) and on the right a window which will display the fields you select for your user-view (if the window is not empty now, clear it by clicking the "Remove All" button):



We will start by cutting down the identification section. Of course we want to know which substance the displayed data belong to, but we do not need all the information given in the record. Open the Characterization folder and the next folder Identification. For our user-view we will select the Gmelin Registry Number, component molecular formula, chemical name, type of substance and CAS Registry Number. You do not have to copy every single field name separately to the right-hand window, but you can click all desired fields and paste them to the right-hand window in one step by clicking the "Copy" button.



Next we will select some data: We are interested in information about NMR-spectra and chemical reactions. If you do not know in which section you can find NMR-data use the "Find" button and type in NMR. The highlight will jump to the entry NMR spectroscopy, which is a folder containing several fields. You can select the whole folder by a single mouse click at the header of the folder:



On clicking "Copy" all fields within this folder will be copied to the right-hand window. The folder Chemical reactions can be transferred to the user-view window in the same way.

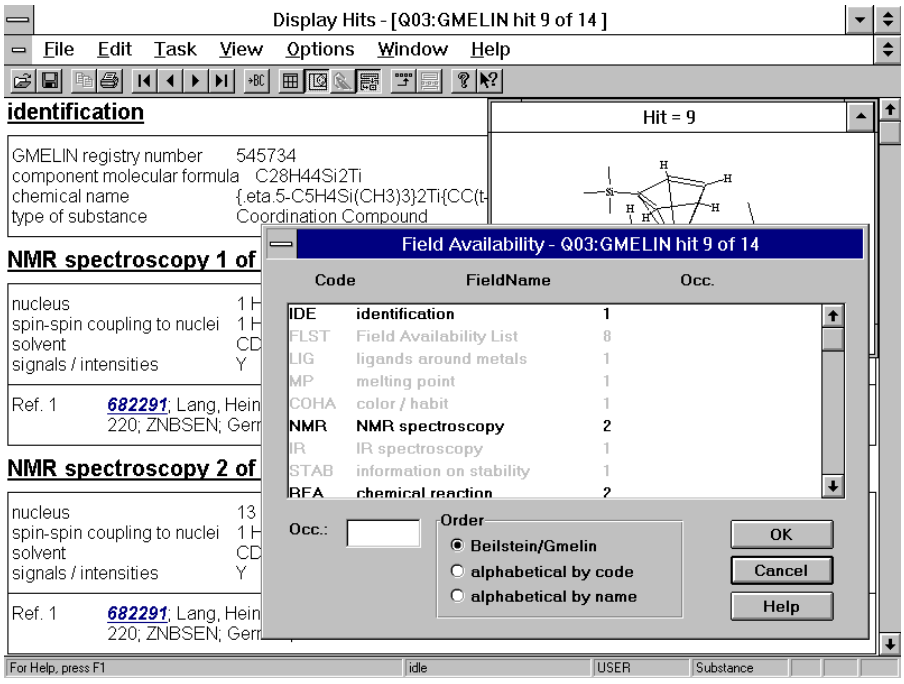
Last but not least we will copy the whole folder "Bibliographic Information". This is necessary to display the references for all the data previously chosen. Of course you could also select only some fields of the bibliographic information, but as it does not need much space in the display anyway, we will not bother to select the most important fields.

Whenever you define a new user-view do not forget to select the bibliographic information in any case when you want to see the references to the original literature.

The definitions of the user-view remain intact until you define a new user-view. Additionally you can save different user-views within the "Define User-View" dialog box and load them again at a later time.

If you click OK the definition of the user-view is finished, but it is not yet active. Select it now from the menu "View".

Now you have reduced the record as you prefer. If you open the "Field Availability" window you will notice that all the hidden fields (which are available for this substance, but not shown in the user-view) are "greyed".

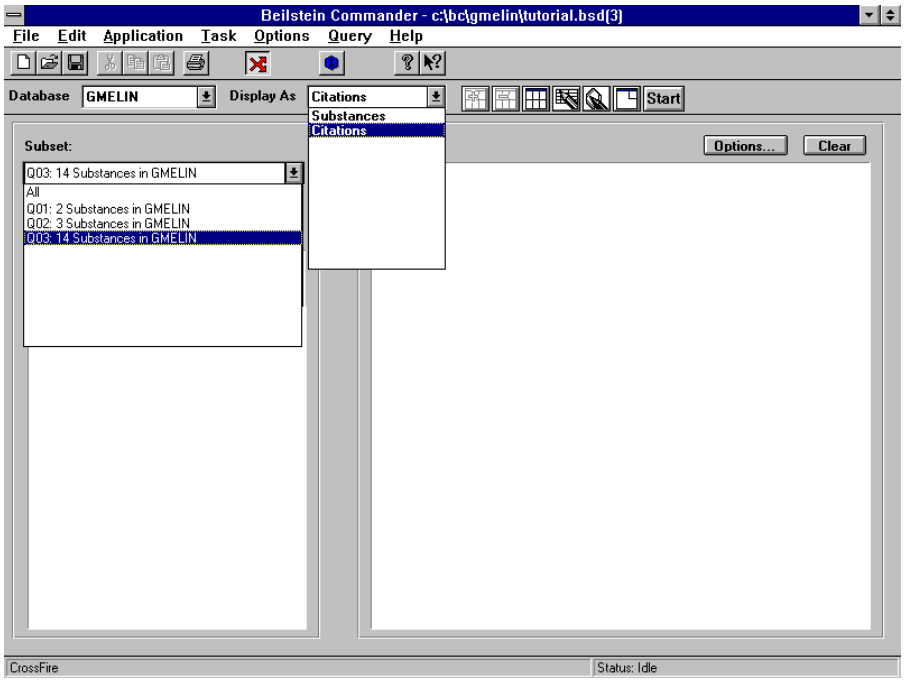


It is important to notice that the user-view applies to printing, i. e. only the data displayed are eventually printed.

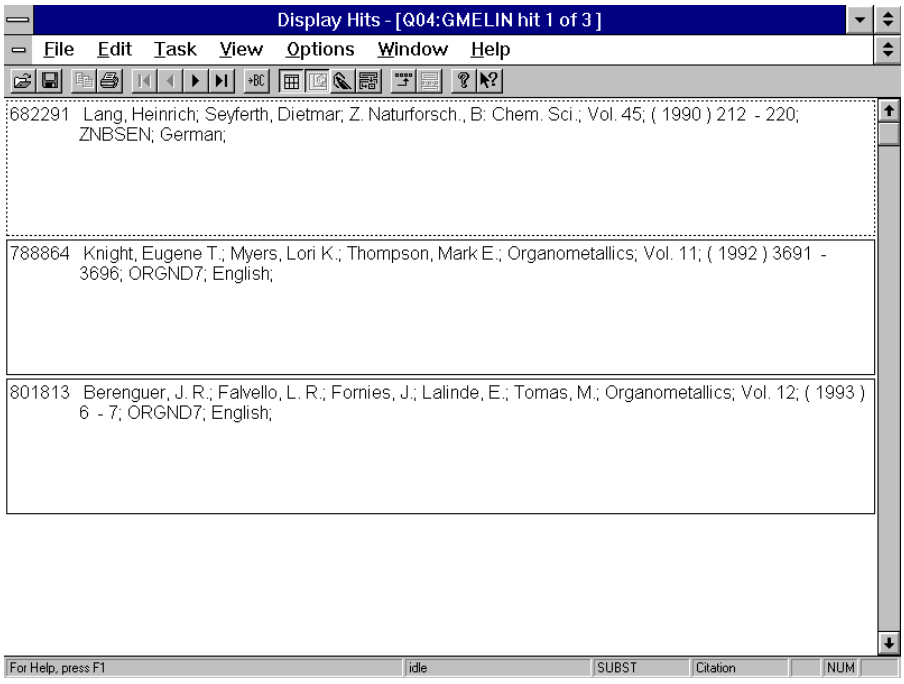
4.6 Viewing the Current Hitset in a Different Context

Let us now explore another feature of the CrossFire system which allows you to display a hitset in a different context. Switch back to the Commander. Look at the context listbox which is at the moment set to Substances. Usually the context is selected automatically by the kind of query you submit: if you do a structure query or a search for a special physical property it will be set to substances; but if you look for an author or a certain journal it will be set to citation mode (you have already seen this kind of display when employing the hyperlink to a reference).

If you want to look up the original literature of the substances of the current hitset you can of course get all the necessary information from the references of the corresponding data, but there is a more elegant way which provides you with a list of all references within all hits of the hitset. To get this you do not have to submit the whole query again, but you can choose the current hitset as a subset and submit an "empty" query with the context changed to "Citation":



As an answer to this query you get three citations; i. e. all the information on the 14 substances of the former hitset, which is collected within the Gmelin database, is contained in only three publications. Switch to the Display Hits. You will see the detailed data of the citation as you have seen before using the hyperlink. But right now you are not interested in the detailed information; you only want to see and perhaps print a list with all references. You will get such a list, when you change to Short Display which can be chosen in the citation context, too. Now you have exactly the handy list you wanted and can take it to the library:



Note that the number of citations shown on the screen in the short display cannot be changed (as you can define the number of structures in the substance context). You will always get four citations at a time.

Now you have met almost all the features of the Display Hits and learned to adjust it to your convenience. This will also help you in viewing the results of the fact queries we will do in the next chapters.

Substructure Search Using the Fact Editor

5

5 Substructure Search Using the Fact Editor

What you will learn:

- ***Using ligand descriptions to search for substances***
- ***Combining fact queries with different operators***
- ***Combining former hitsets with further fact queries***
- ***Using the Hit only view in the display***

5.1 Ligand Classifications

After we have done a substructure query for an organometallic compound by explicitly drawing a structure in the last chapter we will now learn how to formulate similar queries in the Fact Editor using the ligand descriptions within the identification section.

In this example we want to look for a binuclear coordination compound containing either Rhodium or Ruthenium and at least one benzene ligand (any substituents allowed) and one or more Carbonyl ligand(s). Furthermore a halogen ligand should not be present.

First of all a short introduction on the Gmelin ligand classification which can also be found in the Gmelin Reference Guide (Help menu within the Fact Editor)

Ligand Formula (LIGFO)

The search field ligand formula contains the ligand codes according to the GMELIN Ligand Search System which describes the elements (in groups) and the number of the atoms that are connected to the central metal atom(s). A ligand formula is given for all structured compounds or fragments which contain at least one metal atom and at least one coordinating ligand atom of the classes "L, A, D, Q, X" or a special ligand.

The codes include the following elements:

Code	Atom
L	C
A	B, Si, Ge
D	N, P, As, Sb
Q	O, S, Se, Te

X	H, F, Cl, Br, I, At

The denticity of a ligand is put in parentheses in front of the ligand code. If there is a combination of the codes, they are ordered alphabetically.

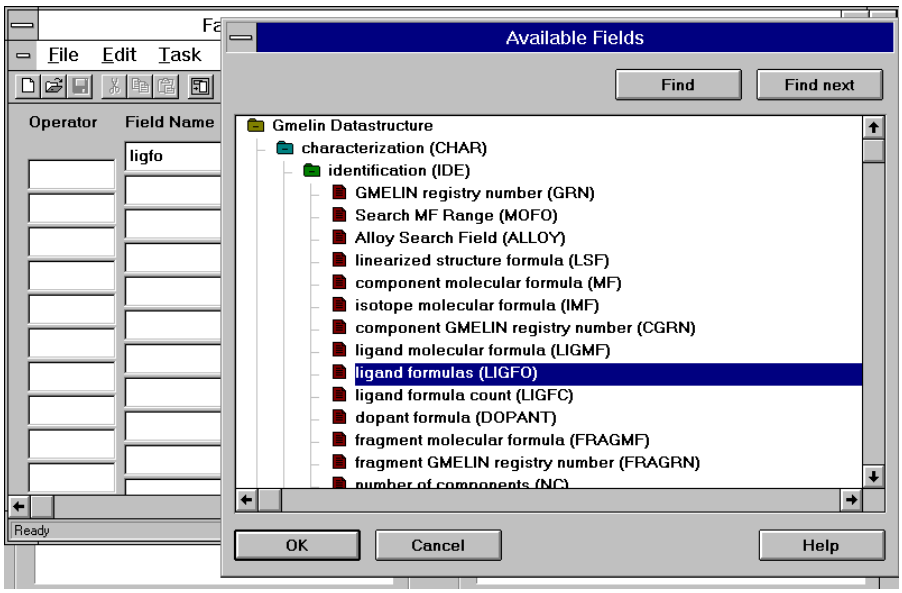
Furthermore there are the following "special" ligands:

CO, CS, CN, CNS, CNO, CNR (the substituent R has no further bond to the metal atoms)

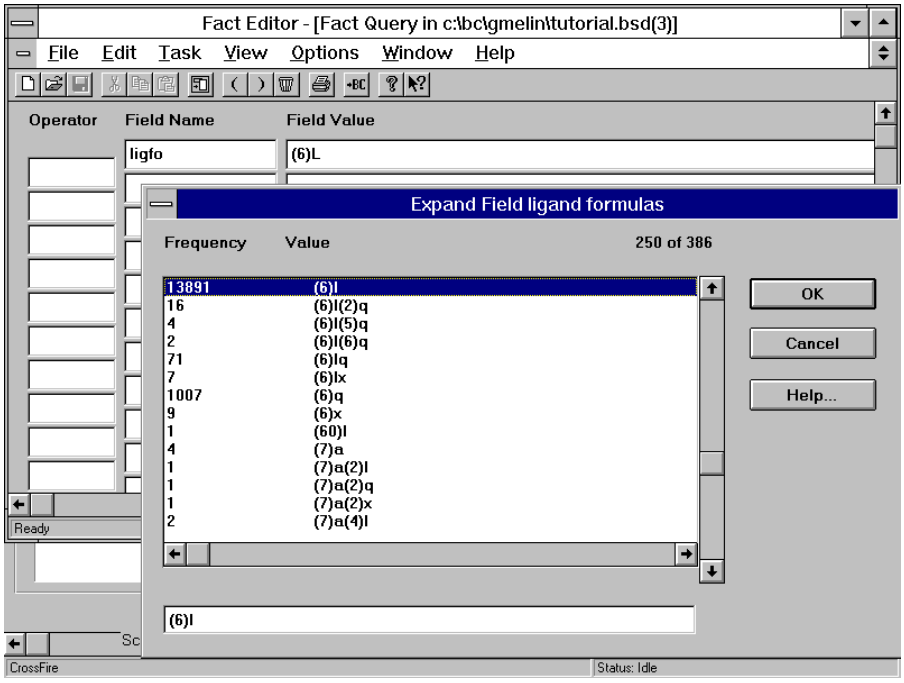
This ligand classification allows you to search for certain types of ligands in a generalized way without having to draw complicated substructures.

As you can see from the table a benzene ligand will have the notation (6)L, 6 designating the denticity and L describing the kind of atoms bonded to the central metal atom. A halogen substituent (which we want to exclude in this query) is represented by an X; in cases of a ligand with only one bond to the metal no number is given for the denticity. Note that other atoms can be present as substituents of the ligand which are not directly bonded to the central atom. In our example this means that although we exclude halogen ligands in our query there may be for example halogen substituted benzene rings as ligands.

After clearing the Structure Editor window in the Commander we will open the Fact Editor. Use the "list values" command to open the Gmelin datastructure. Search for the entry "Ligand Formula" (within the Identification section) and paste it in the first Field Name cell:

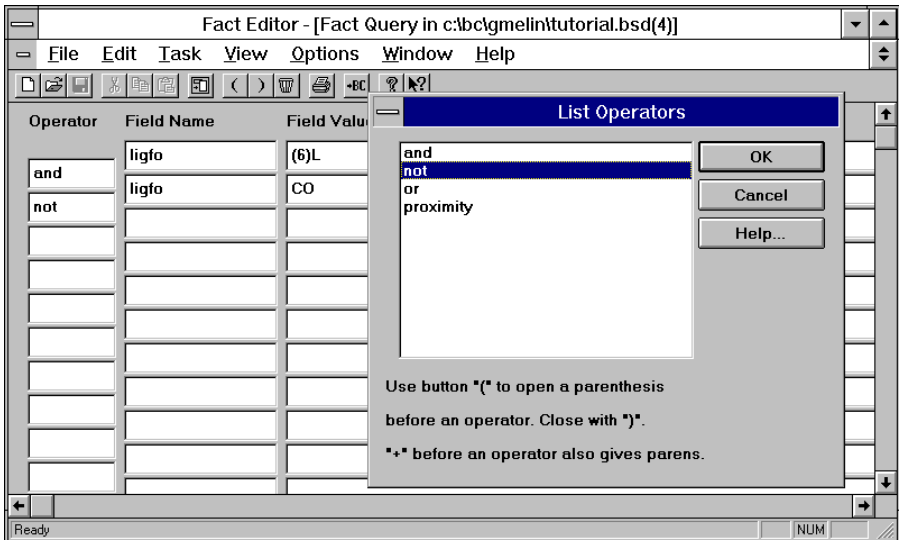


Move to the next column and type in the ligand description for the benzene ring which is (6)L as detailed above. Confirm the correct spelling by the "list values" command:



Move to the operator column and type in "and", move to the next Field Name cell and type in the code for "Ligand Formula" (ligfo) and type in "CO" to the corresponding "Field Value" cell.

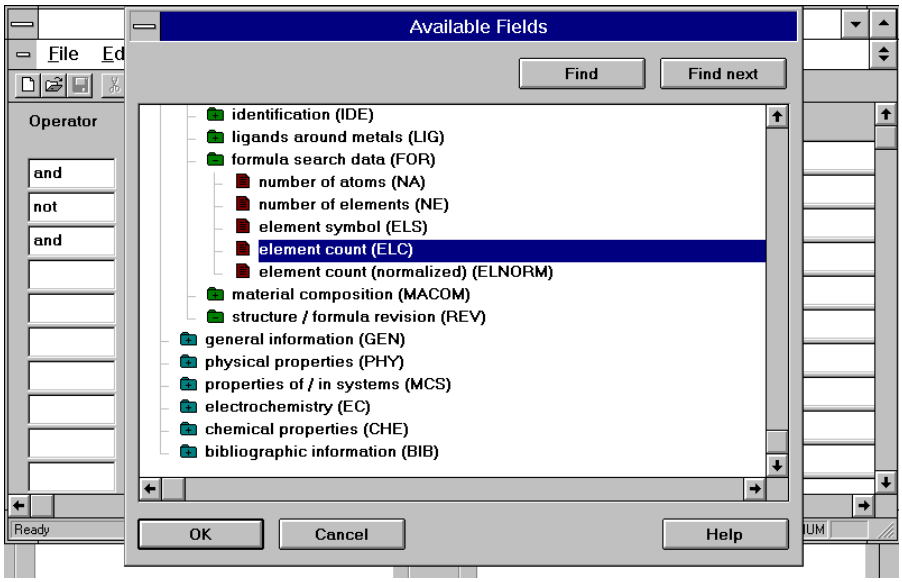
The next step is to exclude any halogen ligand; therefore we have to use the operator "not" to add this part of the fact query. You can see a list of all available operators by using the "list values" command. Make sure the cursor is placed in the Operators column. You can pick the "not" operator from there.



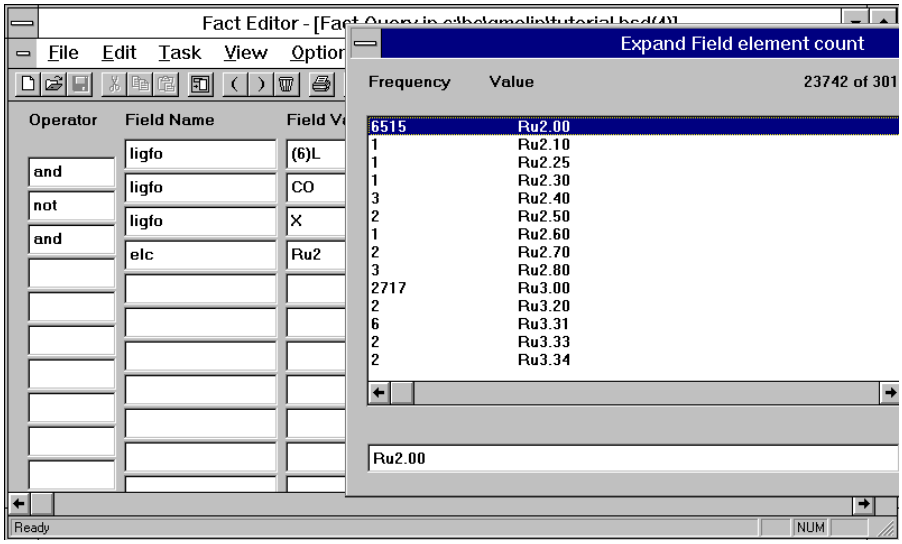
The field code is again "ligfo", and the field value is "X" for the halogen ligand. Now we have finished defining the ligands and we are now going to specify the metal atoms.

5.2 Further Formula Search Data

Within the characterization section there are two fields in which you can search for elements. The first is named "Element Symbol" where you can search for the occurrence of an element in the substance without specifying the number. The second is named "Element Counts" where you can search for any element combined with the number. In this case we want to search for complexes containing either two Rhodium or two Ruthenium atoms, so we will use the element counts. For this kind of search terms you will find a special subsection named "Formula Search Data" within the characterization section.



Select "Element Counts" from the datastructure and move the cursor to the next cell. Type in Ru and without a space add 2. Click the list values button:



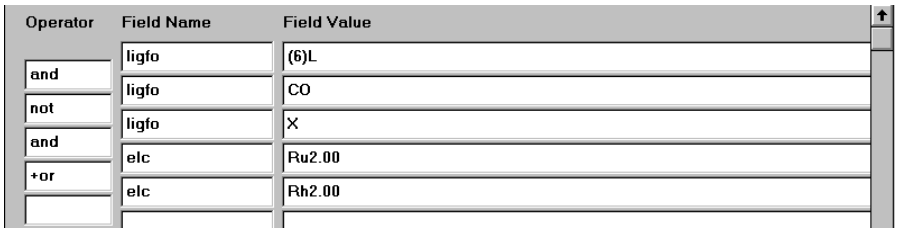
You will notice that all the element counts are listed with two digits, so we will pick the entry "Ru2.00" from the list. The next line has to be combined by "or"; since you know by now how the index for the element counts works you can type in directly "elc" and "Rh2.00".

Now we have specified all requirements for our search, but we have to add one more parameter. The different operators follow a certain order and if you want to

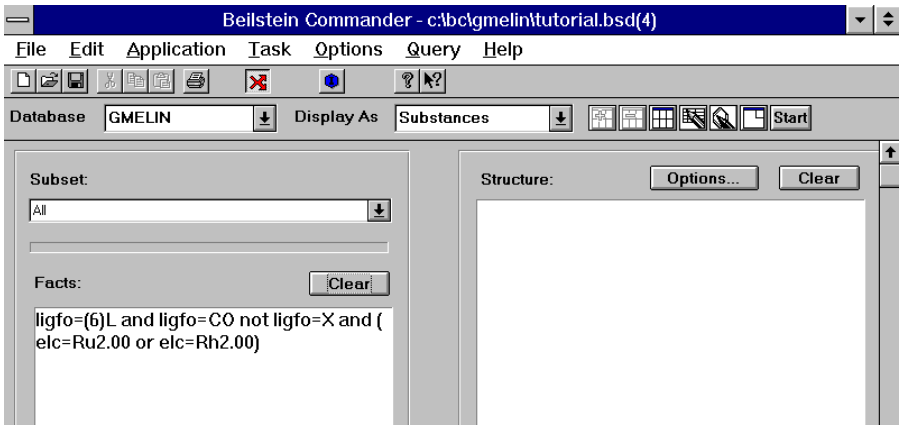
combine the queries in a different order you have to set parentheses. The order of the operators is as follows:

proximity > and, not > or

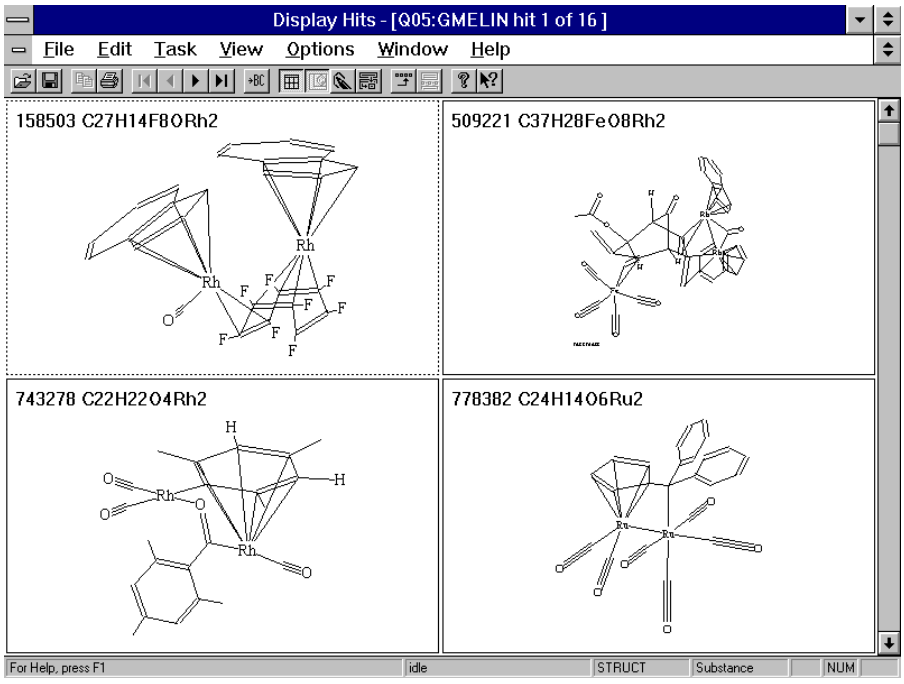
Therefore we need to put the two lines combined by "or" into parentheses. Place the cursor in front of the first "elc" symbol and click on the button for opening the parentheses (( ). Move the cursor to the second "elc" symbol and close the parentheses ( )). You will notice that no parentheses are shown within the Fact Editor, but a "+" appears in front of the "or":



When you are familiar with this convention you may also type in directly a "+or" into the operator cell. If you switch now to the Commander you will notice that in the Fact Editor window the parentheses are shown as you placed them within the query:



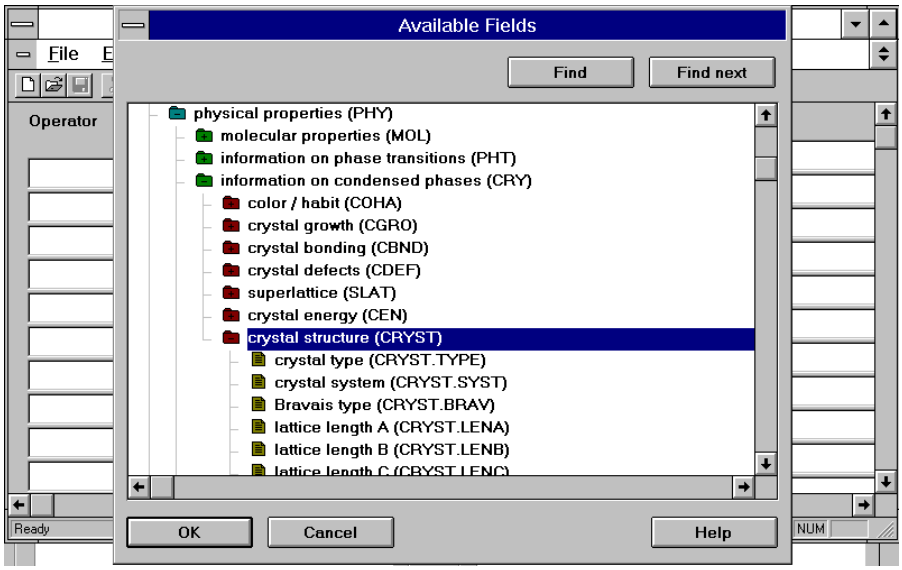
Start the search an have a look at the retrieved hits in the short display:



You will notice several different types of substances but you can also see that all the hits fulfil the criteria of our query. Now perhaps you are interested to know if any of the structures has been assigned by an x-ray analysis. Of course, you could browse all the records for such data, but it is much more convenient to combine the search for crystal structure data with the formerly retrieved hitset. How this works we will see now.

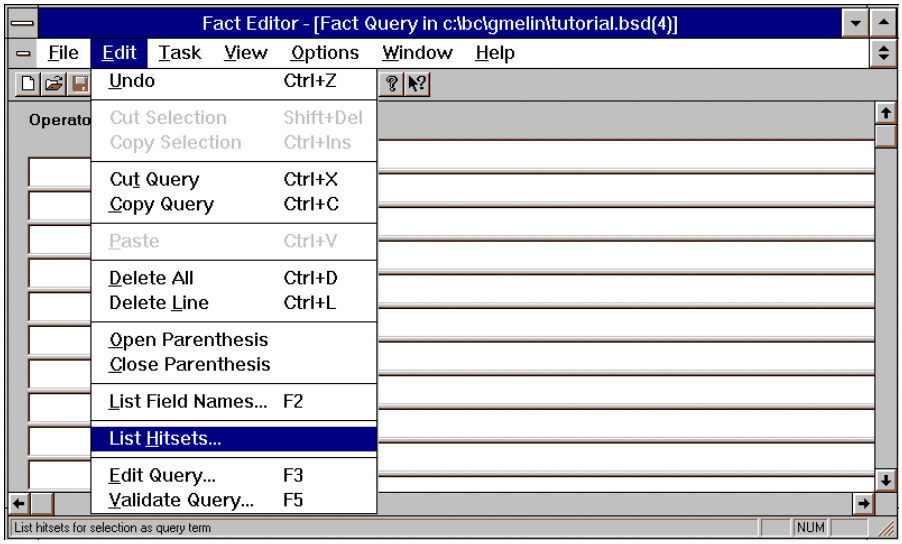
5.3 Combining Former Hitsets with Further Fact Queries

Switch back to the Commander, clear the Fact Editor window and open the Fact Editor. Open the Gmelin datastructure and use the "Find" button to locate data on the crystal structure by typing in "crystal". Select the appropriate field code "cryst":

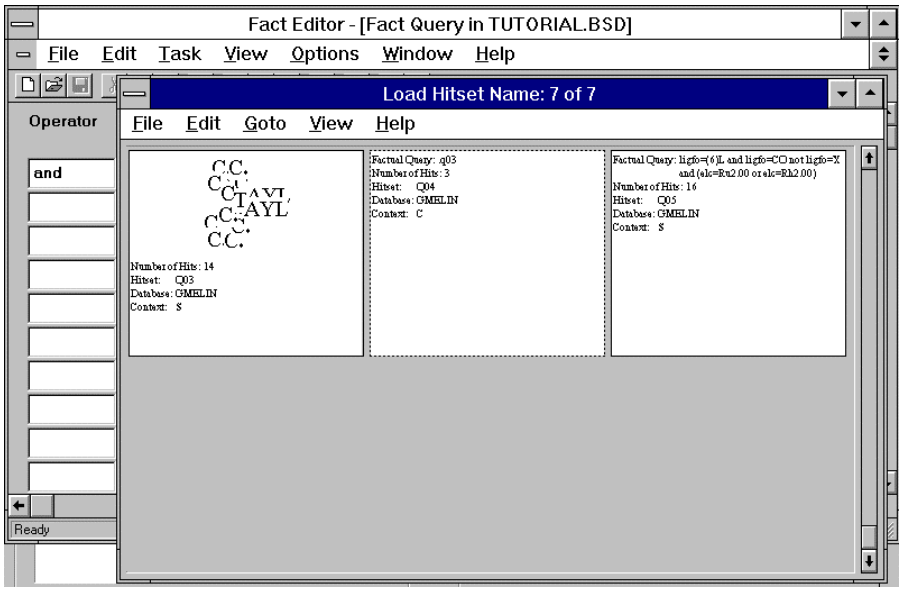


This time we do not fill any value in the "Field Value" cell since we just want to know if there is any information available. This is a so-called "Field Availability Search".

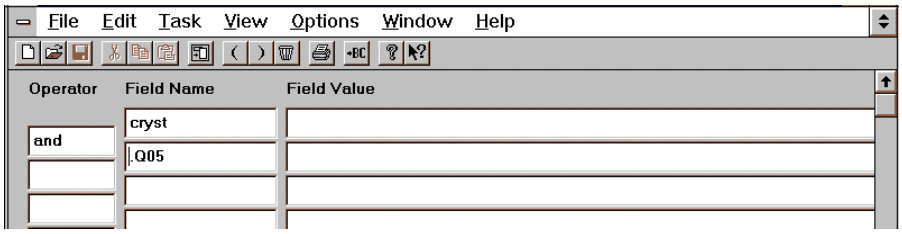
Next we will combine this query with the substances of the formerly retrieved hitset. You may have noticed that all the hitsets are numbered with Q01, Q02 etc. You can view a list of all the hitsets of the current session if you place the cursor in the next Field Name cell:



Choose from the edit menu "List hitsets...": an additional window opens which looks similar to the short display and shows all your submitted queries with some more information such as the context of the search, from which database the hits were retrieved, the number of hits and of course the number of the hitset:



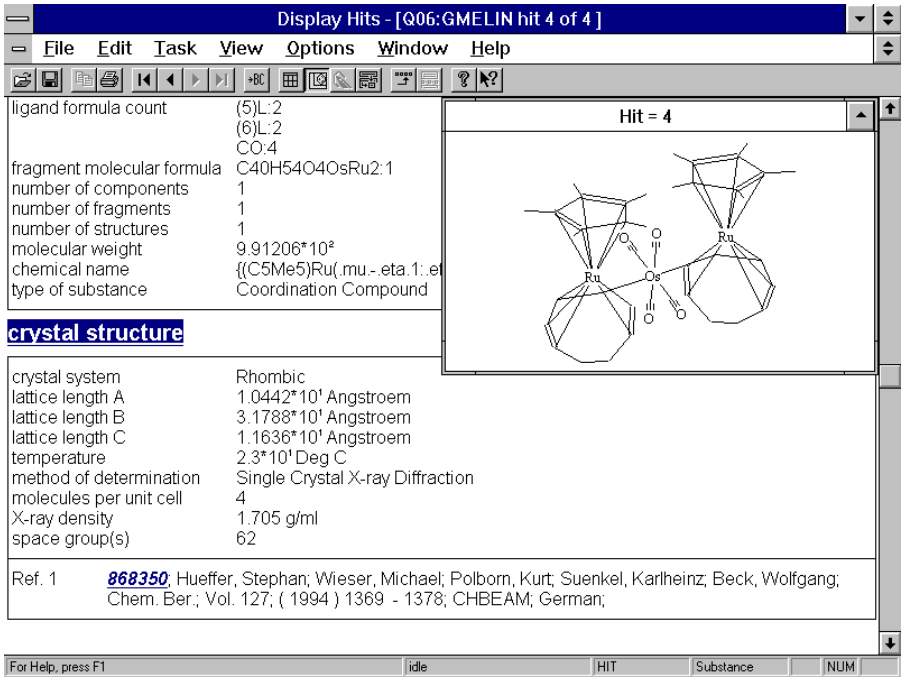
If you click now the last entry (the hitset we had retrieved in the beginning of this chapter) the selected number will be placed in the Field Name cell (in this case it is the number Q05, yours might be different). If you remember the number of the hitset you can also type it in directly: be sure to type a dot in front of the Q-number to indicate to the system that you refer to a hitset:



After switching back to the Commander and starting the search you will get four hits.

5.4 The Hit Only Display

Let us have a look at them. Within the Display Hits select the Full Display again (either from the View menu or by the "Change Full/Short" button). Open the View menu again an select this time "Hit only". By this you set a filter to the display which will now only display those data you have searched for:



With this "Hit only" mode you can quickly view the hits without having to browse the whole record to find the data you are interested in.

After having finished this fourth chapter of the tutorial you are now familiar with all the important features of Structure and Fact Editor as well as the Display Hits. In the next chapters you will now learn more about the search facilities the CrossFire system offers to you. It will show you that there are more possibilities to gather information than by looking for particular substances.

Searching Substances Possessing a Certain Property

6

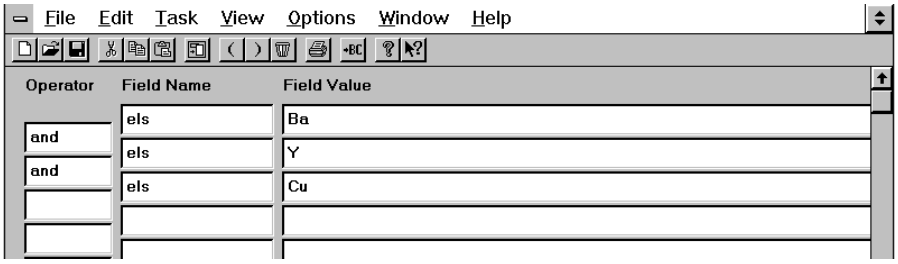
6 Searching Substances Possessing a Certain Property

What you will learn:

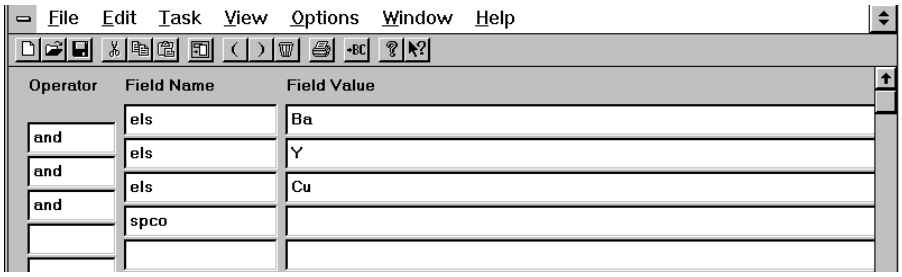
- ***Using further possibilities to formulate queries in the fact editor***
- ***Searching numerical fields with ranges***

In this example we are going to look for superconducting substances which contain Barium, Yttrium and Copper with a special interest in the critical temperature of the superconductivity.

Again we will use the Fact Editor to formulate the query. You may remember from the last chapter the subsection "Formula Search Data" within the characterization data. This time we will use the field "Element Symbol" since we are not interested in a certain composition of atoms. All those queries will be combined by "and":



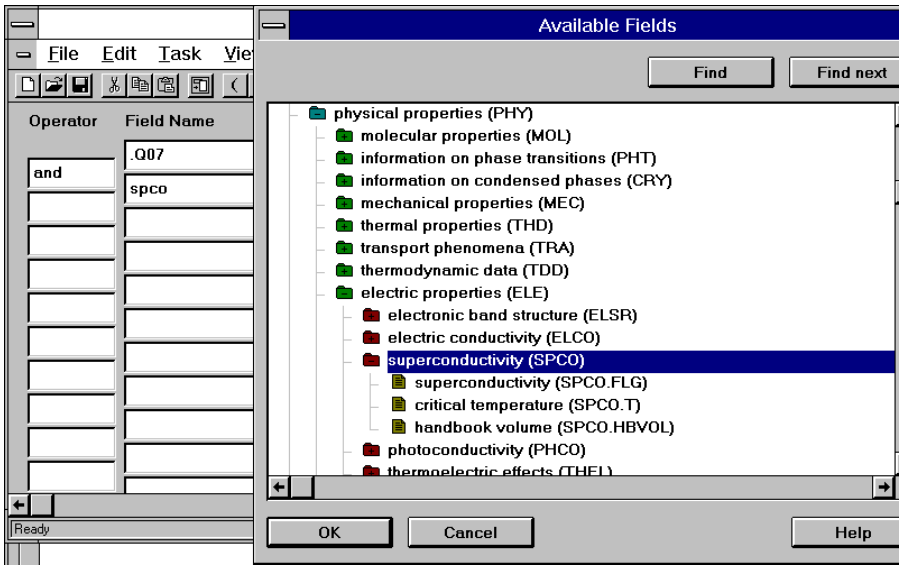
Then we will add the query for superconductivity; use the "list values" command and the "Find" button to locate the appropriate field. At first we will do a "Field Availability Search" without defining a value:



Switch back to the Commander and start the search. You will get 3029 hits. In a case like this it will be necessary to cut down the hitset and specify the query (unless you would like to work through all the hits). You do not have to open this large hitset at all, just click the "Close" button; the hitset is saved on the server even though you did not look at it. Switch to the Fact Editor again.

If you have not cleared the Fact Editor window within the Commander beforehand you can use the shortcut "Ctrl + D" to delete all lines in the Fact Editor table. For deleting just one line use "Ctrl + L"; make sure the cursor is placed in the line you want to delete!

As in the last example we do not have to perform the search again, but we can combine the former hitset with an additional query. You will notice within the datastructure some more attributes assigned to superconductivity.

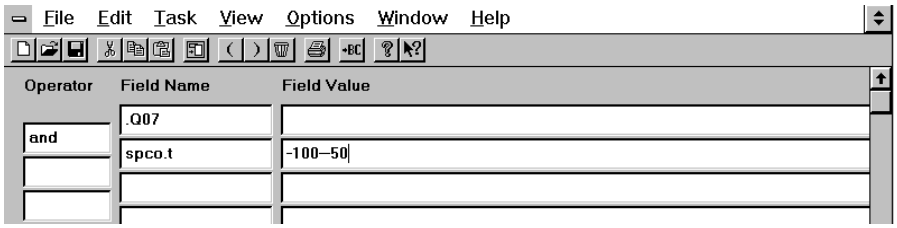


It should be useful to formulate a query for the critical temperature to retrieve the most interesting substances. Type into the "Field Name" cell the number of the

former hitset (.Q07 or whatever number your hitset has got) and use the field code "spco.t" for the critical temperature.

When you look for numerical values like this you can fill in ranges as well as use the operators ">" and "<" (Note that "<" always means "≤" and consequently ">" means "≥").

For the specification of this query let us use the range from –100 to –50 °C, which looks like:



Please keep in mind that the operators "-", "<", ">" can only be used for numerically indexed fields. Use the Gmelin Reference Guide (Help menu within the Fact Editor) to learn about the type of index of a certain field.

Additionally it is necessary for most of the numerical fields to know which unit is used for physical properties within the database. They can also be found in the Gmelin Reference Guide but they are given on top of the "Expand Field Values" table, too.

After performing the search you will get two hits, which is a much more handy number to look at. You may decide by yourself which is the best way to view the records; you can use either the "Hit only" mode or you could define a user-view, including besides the IDE-data and the superconductivity perhaps the chemical reactions, if you are interested in the preparation of those compounds.

Furthermore, if you think that the hitset has been cut down too much by the more restricted query for the critical temperature and you feel to have retrieved too little hits, you might submit a new query with a wider temperature range or an altogether other kind of further specification.

When you look at the hits you will also notice the lack of structure drawings. Compounds like these ceramic materials belong to the classes of substances in Gmelin which do not posses a structure. Keep this in mind when you formulate queries; those substances can only be retrieved by factual queries.

Substructure Search with User-defined Atom Lists and Generic Groups Combined with a Factual Query

7

7 Substructure Search with User-defined Atom Lists and Generic Groups Combined with a Factual Query

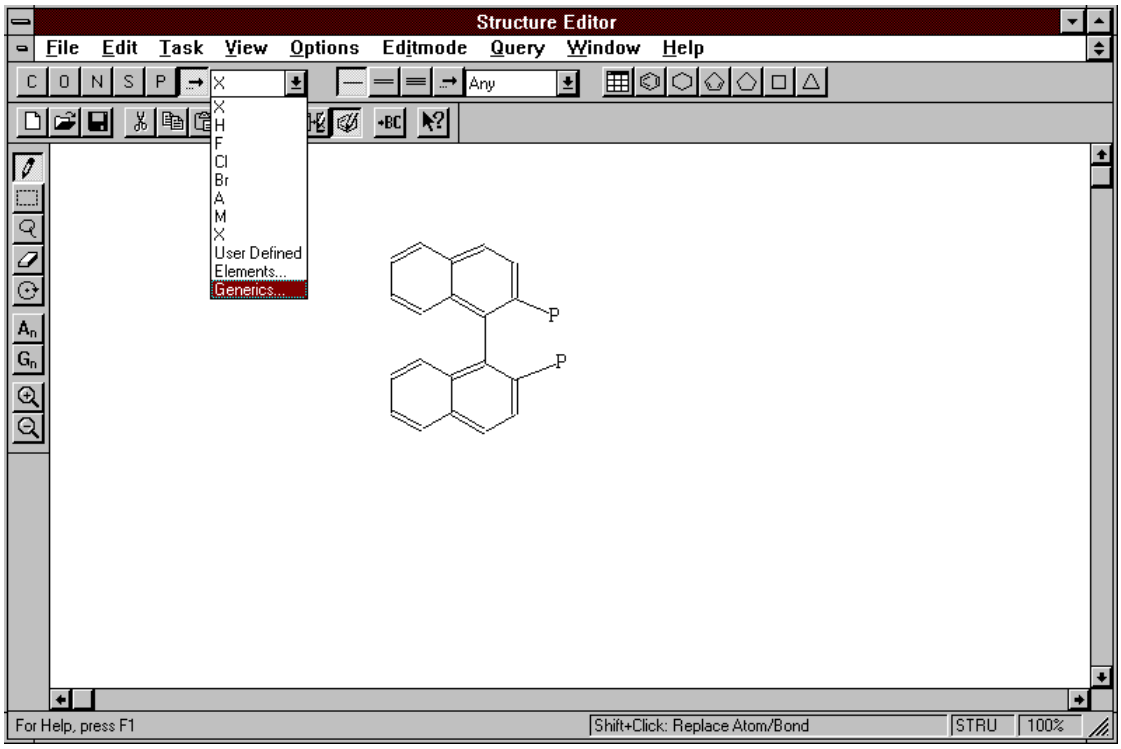
What you will learn:

- ***Employing user-defined atom lists***
- ***Employing user-defined generic groups (definition and setting of attachment points)***
- ***Combination of a substructure and a factual query***

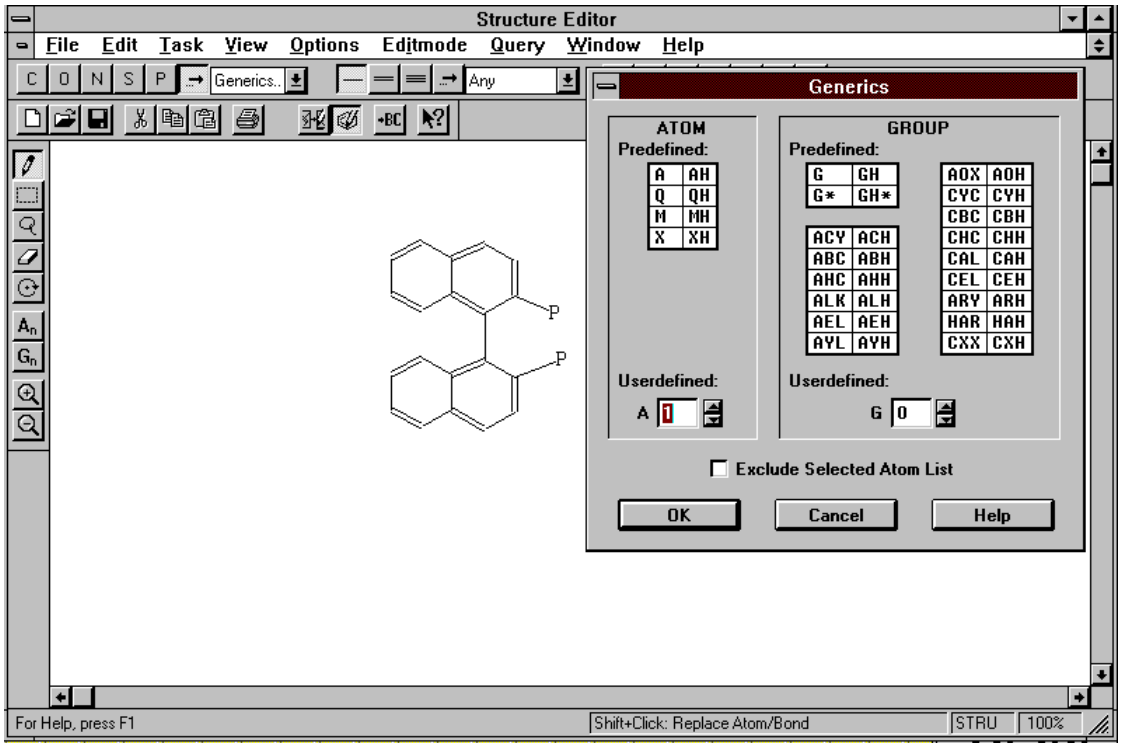
We are searching for metal complexes with a 2,2'-diarylphosphino-1,1'-bisnaphthyl-ligand with respect to their use as a catalyst. As aryl substituents we want to allow phenyl, 3- or 4-alkylphenyl or 3,5-dialkylphenyl groups and as the metal centre we choose all the precious metals Ruthenium, Rhodium, Palladium, Osmium, Iridium and Platinum.

7.1 Defining User-defined Atom Lists

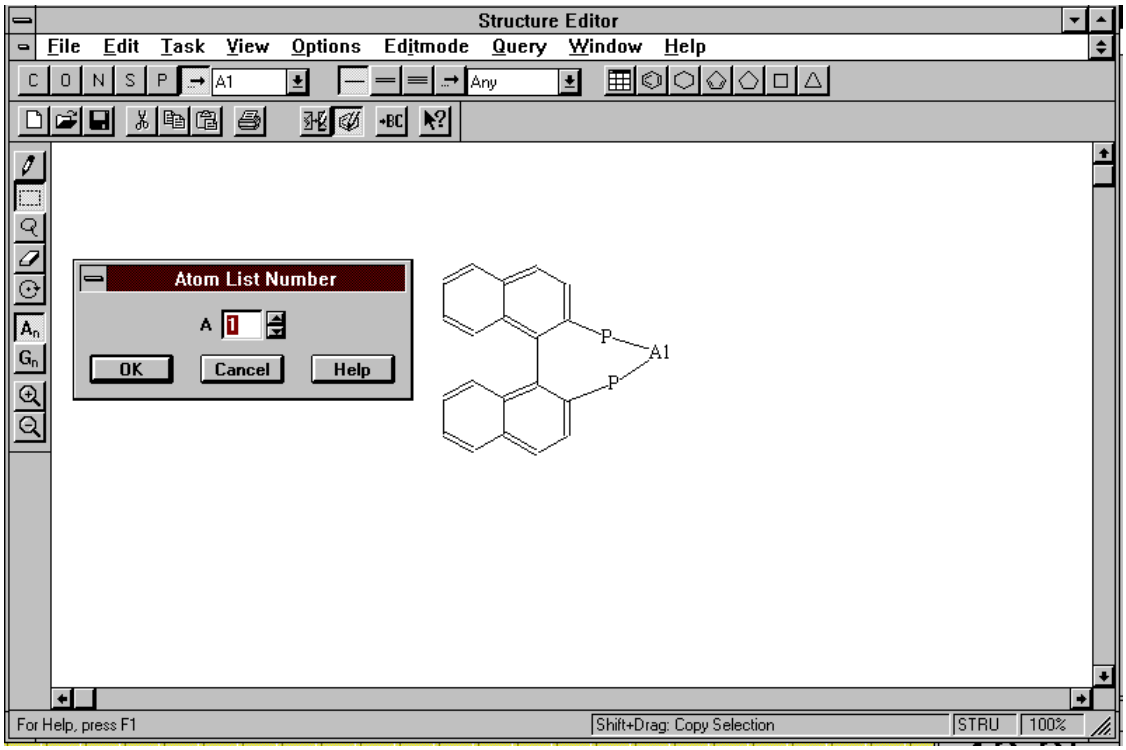
Draw the 1,1-bisnaphthyl group using ring templates and the "Shift and Drag" function in the select mode. Afterwards add the two phosphorus atoms in the 2-positions. Now we will add the central metal atom as a user-defined atom list. Choose from the atom list box "Generics...":



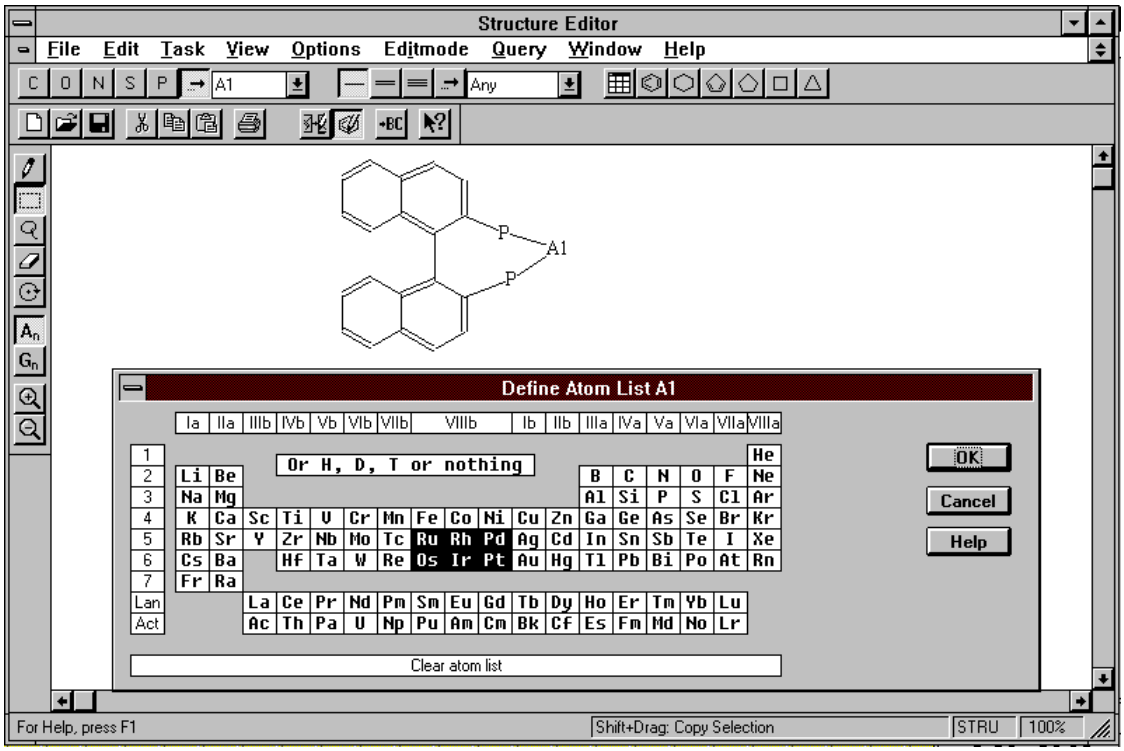
Then select a number for the atom list you are about to define (in this case choose 1) and click OK.



Now you can put the A1 symbol in the parent structure; remember to draw the coordinative bonds as single bonds. After placing the A1 symbol in the structure drawing you can define the atoms belonging to it by clicking the An-button on the tool bar:

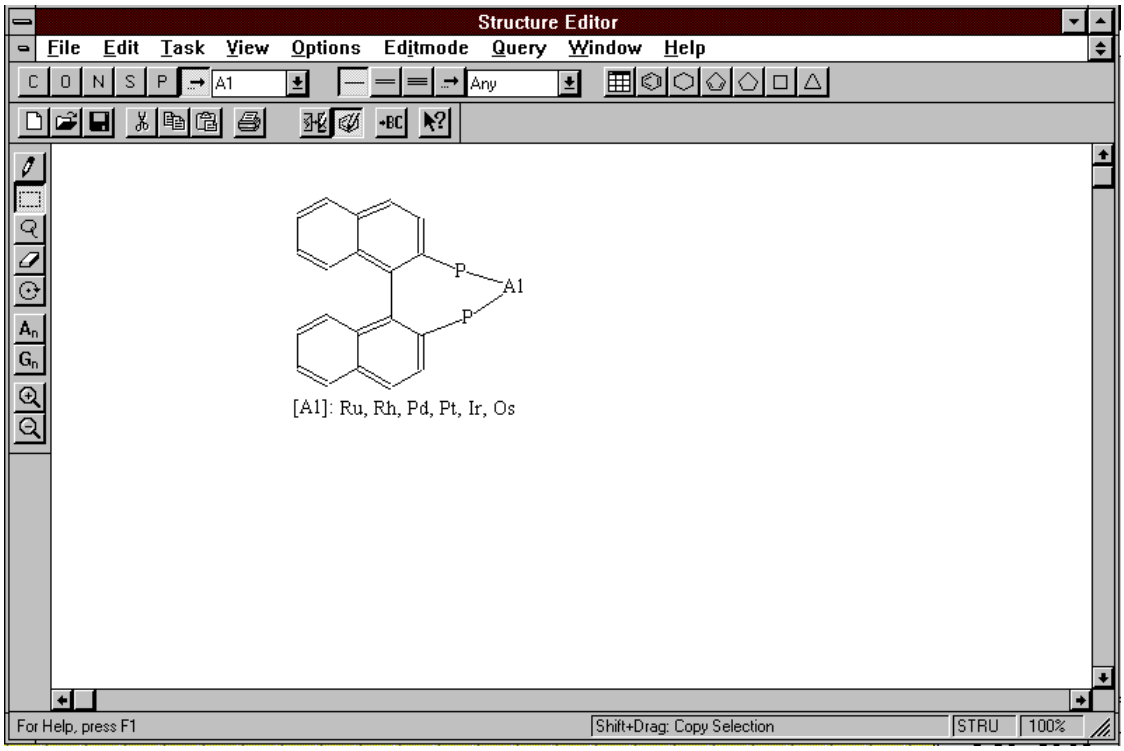


Confirm the chosen number for the atom list by clicking OK and a periodic table will be opened from which you can pick the desired atom symbols:



Note: Do not mix up this periodic table with the one which opens when you choose "Elements...". The "Elements..." table allows you to pick just one symbol whereas you can choose any number and combination of symbols in this "Define Atom List An" table. By a single mouse click you can select a symbol; another click on the same symbol will undo your selection. You can also select a whole row or column by clicking the "header" of it. Last but not least you can delete the whole selection with a click on the "Clear atom list" button.

In this example we choose the symbols Ru, Rh, Pd, Os, Ir and Pt. Click OK and the definition will be written beneath the structure drawing:

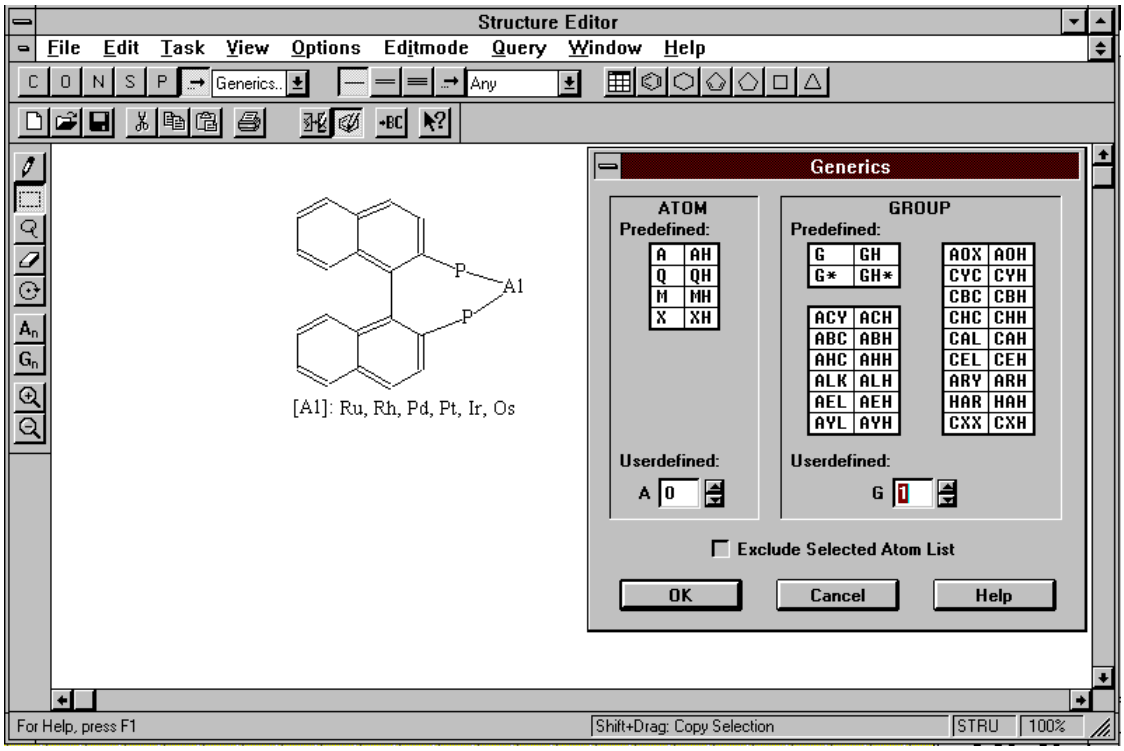


Whenever you define an atom list take care to follow the order described above step by step. Otherwise you will get error messages while trying to define the list. Just to keep those steps in mind a short summary of the procedure:

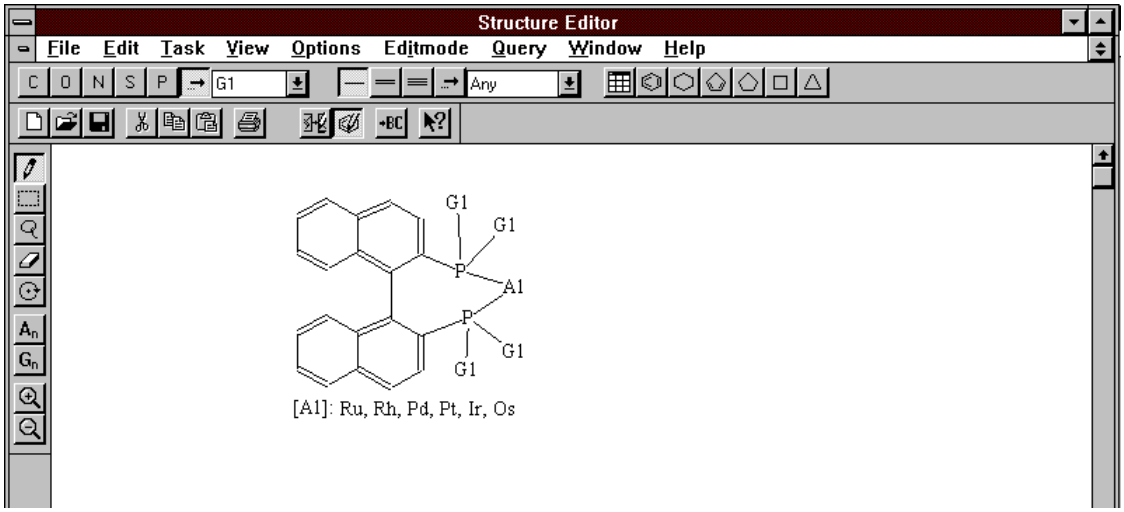
- Choose the A1-symbol and place it in the parent structure
- Click the An-button on the tool bar and choose a number
- Pick the desired elements from the periodic table which opens automatically and click OK
- You can change the atoms of an already existing atom list by clicking the An-button again and choosing the appropriate number; the corresponding periodic table opens and you can change your selection

7.2 Defining User-defined Generic Groups

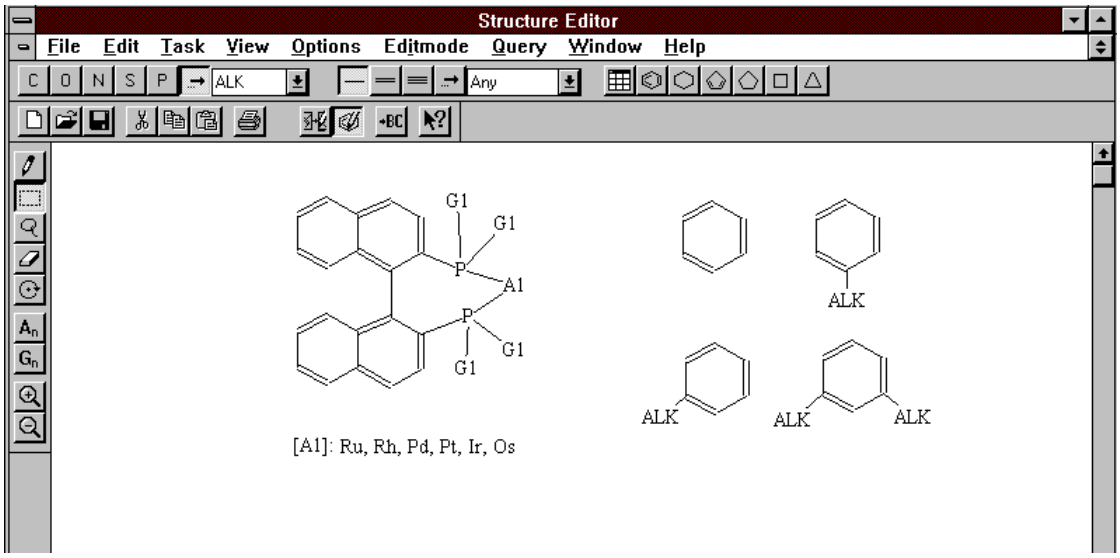
The next step will be the definition of the aryl substituents of the Phosphorus atoms. For that purpose we will employ the user-defined generic group G1. This can also be chosen from the Generics dialog box:



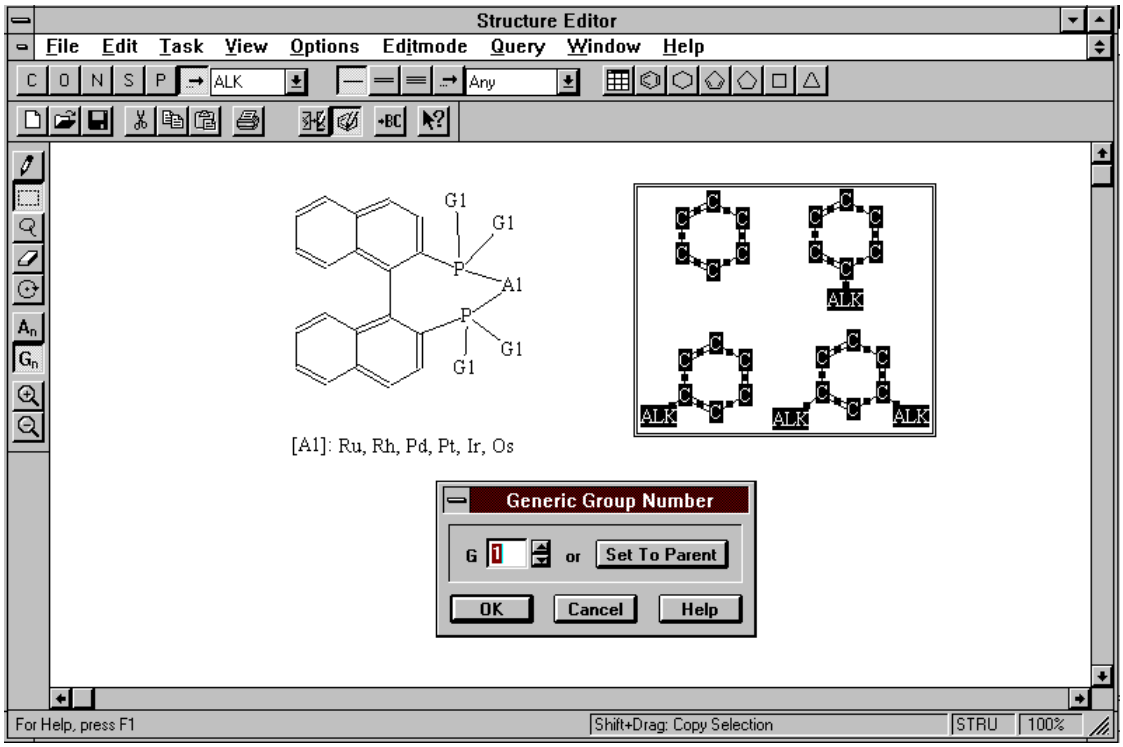
Place the G1 symbol in the structure as substituents to the Phosphorus atom:



Next we will draw the structure fragments which the generic group should contain. Use the pre-defined generic group ALK as substituent for the phenyl groups:

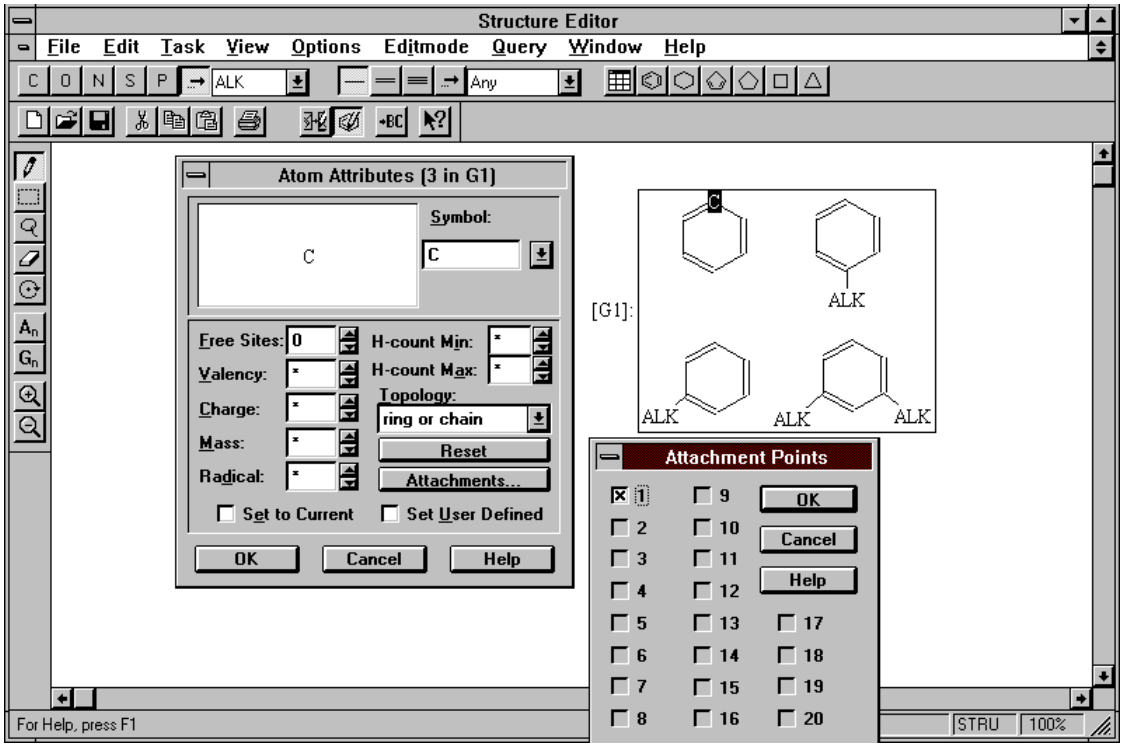


After completing the structure fragments select all of them, click the Gn button on the tool bar and choose the number of the generic group (which will be 1 in this case):

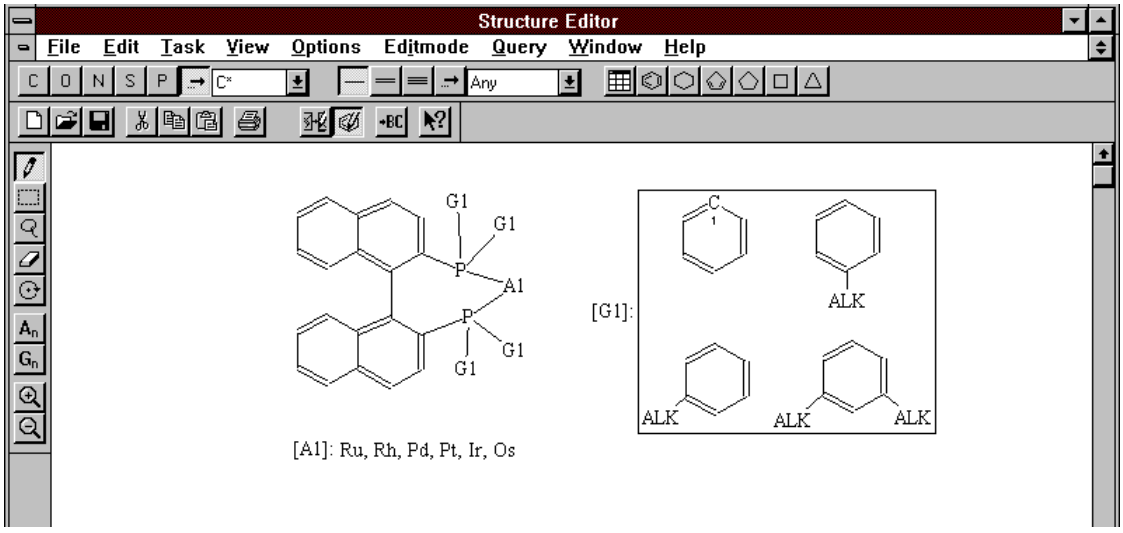


After having defined the fragments as members of a generic group you can set attachment points to indicate by which atom the substituents have to be bonded to the parent structure. Select the edit mode (pencil) and click at the desired atom of

the substituent within the G1 group. An atom dialog box opens where you can click the "Attachments..." button to open an additional dialog box:

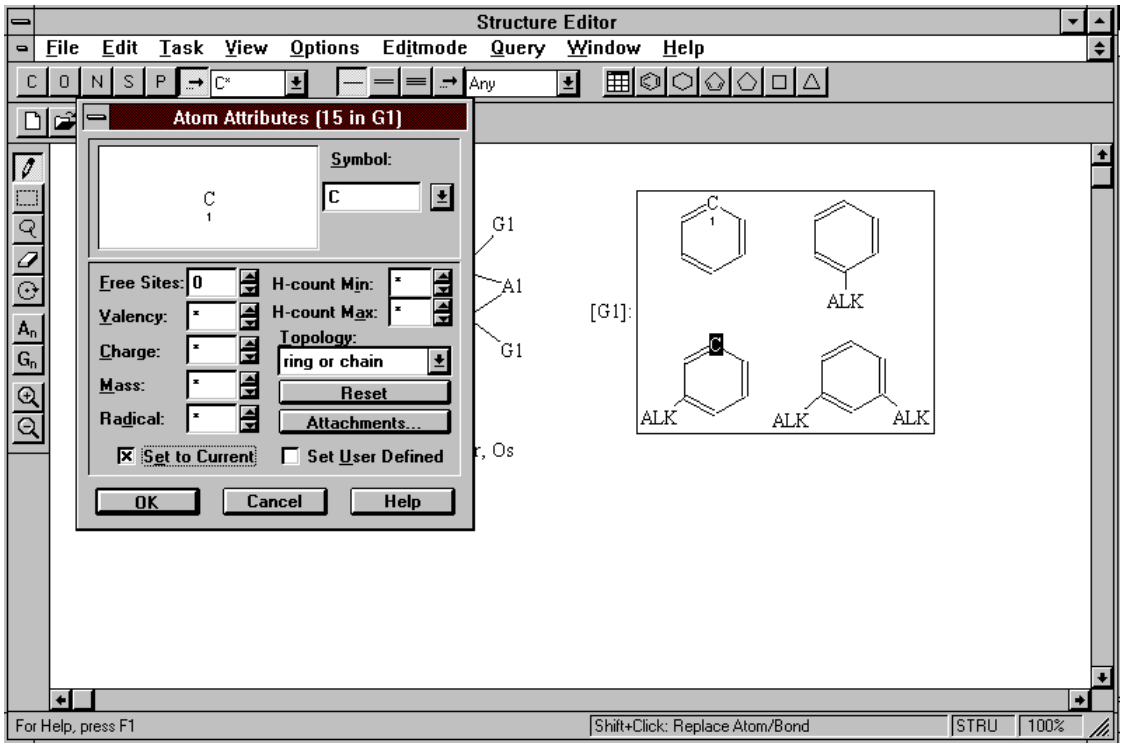


Click 1 and confirm your choice with OK, leave the atom dialog box by clicking OK again. You will notice a small "1" at the carbon atom of the phenyl substituent to indicate the attachment point:

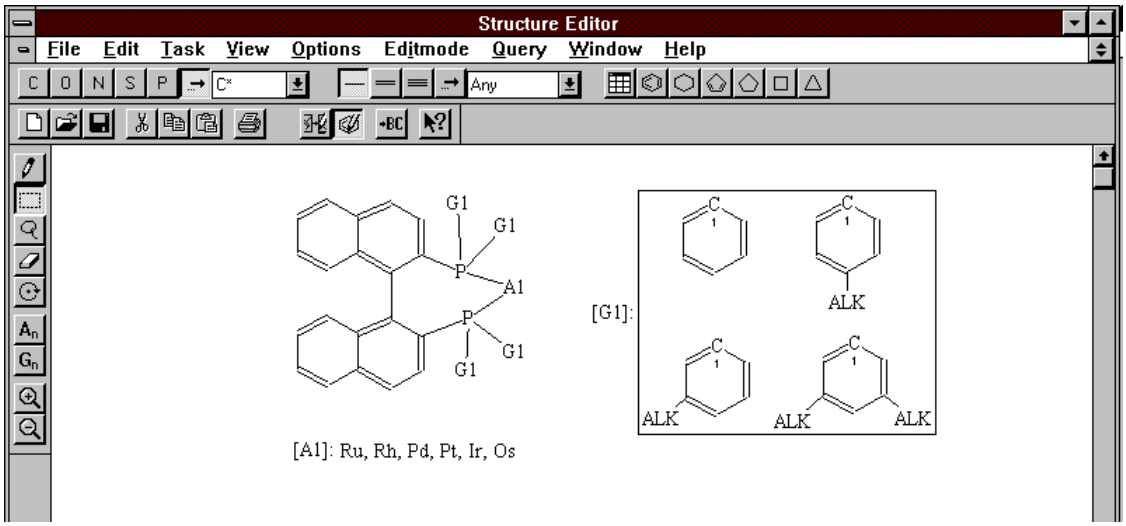


You could now repeat this action for the other phenyl substituents, but there is a faster way to set all necessary attachment points. Click the next carbon atom which will be an attachment points in the second phenyl substituents. Select "Attachments...", click "1" and leave this dialog box with OK as before, but do not

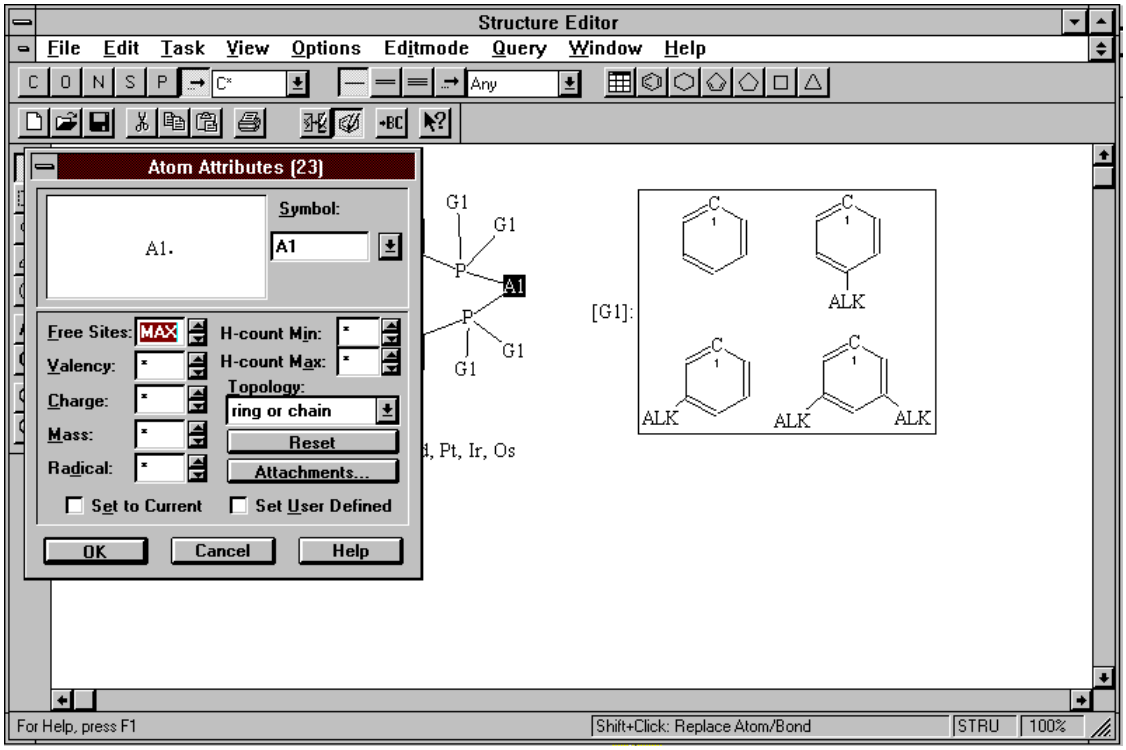
close the atom dialog box yet. At the bottom of the box you will observe the option "Set to current"; if you click this, the just defined atom (with all its attributes) will be placed in the atom list box and will be active until you choose a different atom



Now use the "Shift and Click" function to set attachment points for the next two phenyl substituents: move the mouse cursor to the desired atom, hold down the Shift key and a click with the mouse will replace the carbon by a carbon with an attachment point:



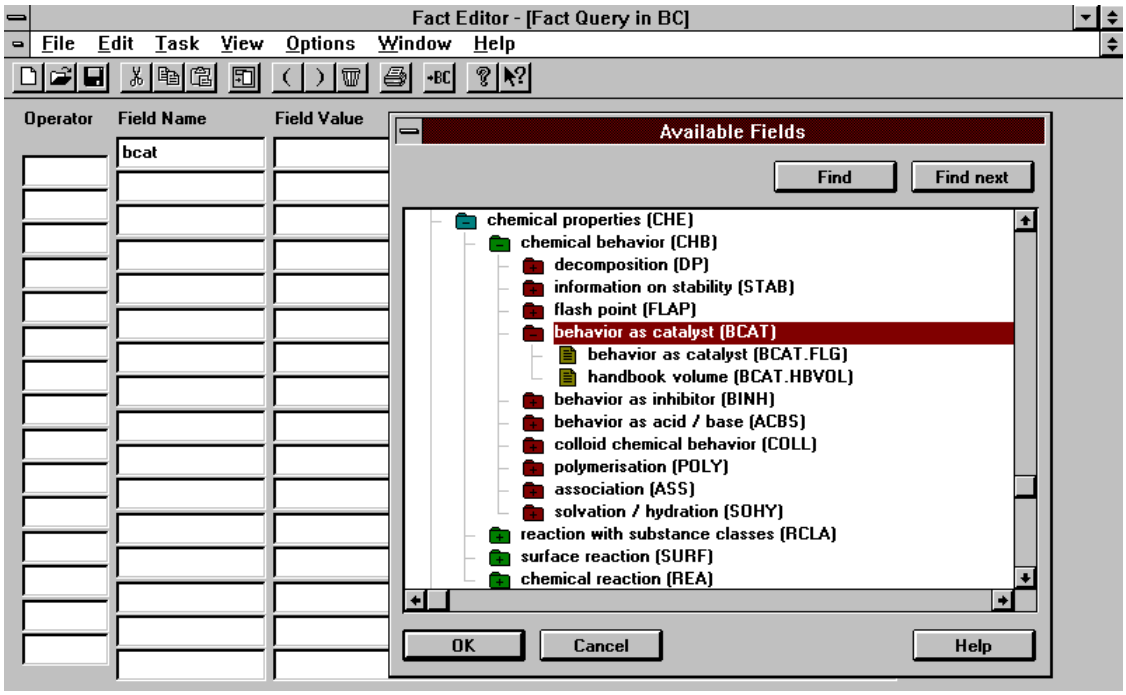
As a last definition for the structure query we will set maximum free sites at the central metal atom represented by the A1 symbol:



After transferring the structure to the Commander we can start to define the factual part of our query.

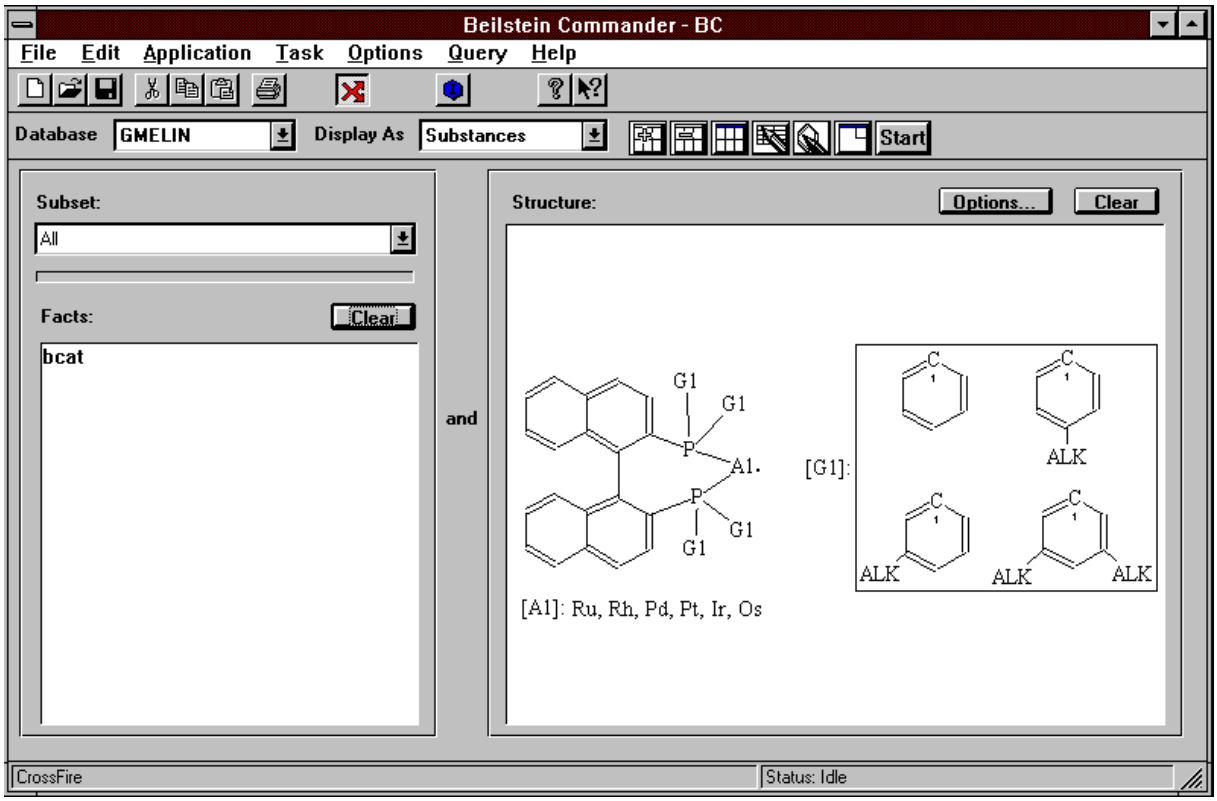
7.3 Combining a structure search with a factual query

Open the Fact Editor and select from the datastructure "Behaviour as a Catalyst":

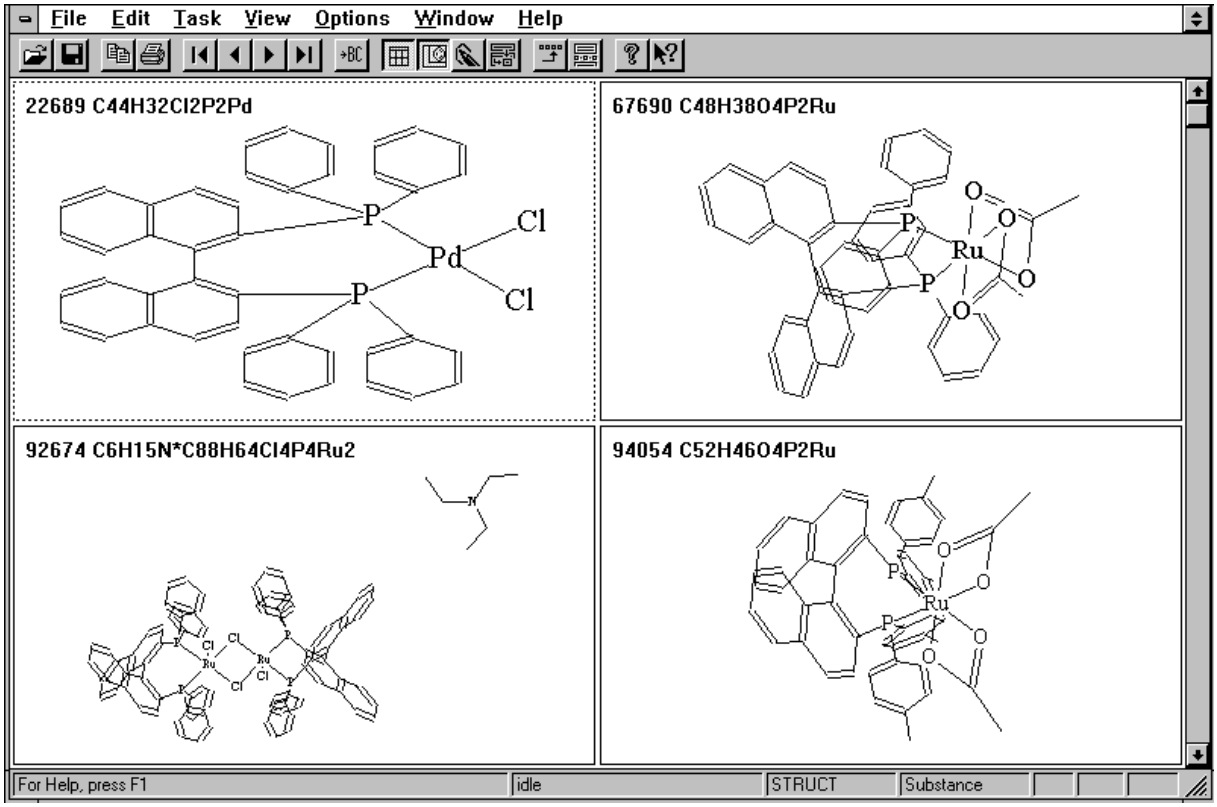


In this case we will again do a field availability search without filling in a field value.

If you switch back to the Commander now you will notice that the two queries in the Structure and Fact Editor windows are automatically combined by an "and":

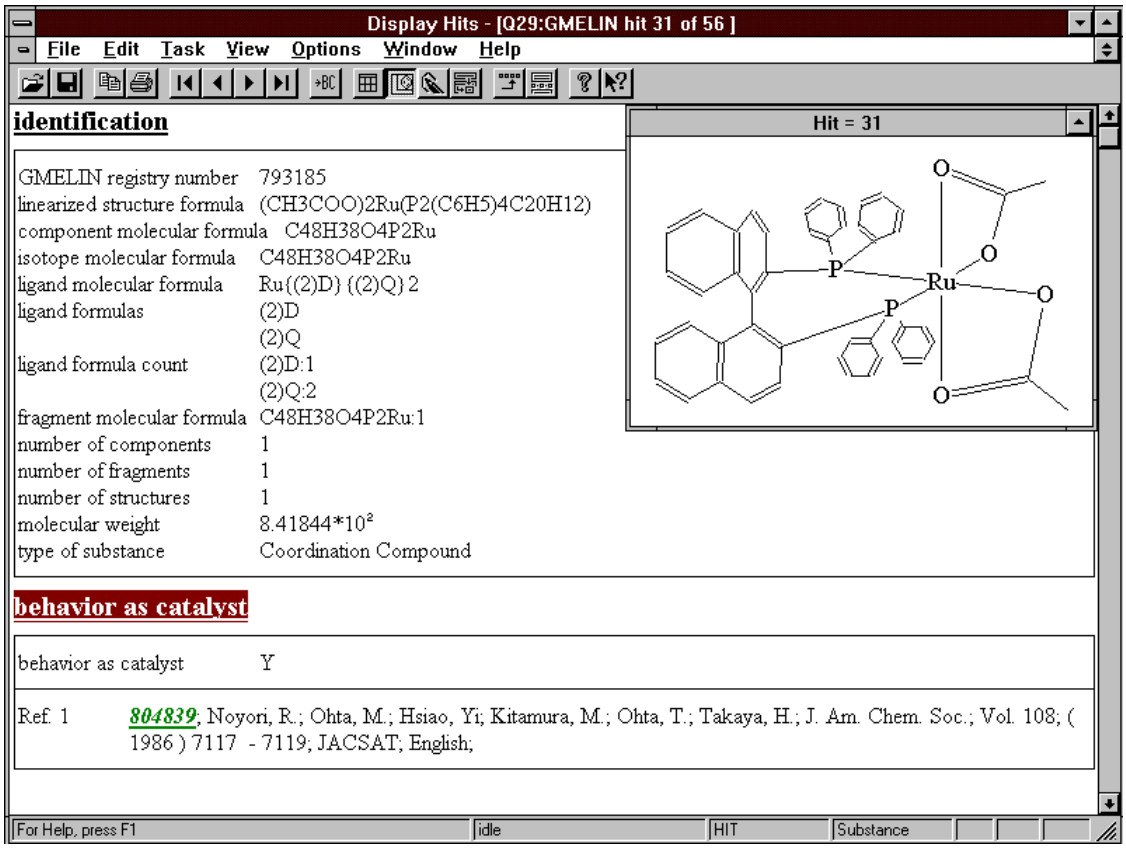


After performing the search you will get 65 hits which can be viewed in the short display:



You will notice that all those structures fulfil the criteria set in the substructure query.

You can also view the hits in the Hits Only display where you can see the entry of the behaviour as a catalyst with the corresponding citation:



In this example you have seen how to formulate a substructure query with user-defined atom lists and generic groups so as to adjust it to your personal needs. It allows you to specify very distinctly the structure you want to search.

Searching for the Preparation of a Substance

8

8 Searching for the Preparation of a Substance

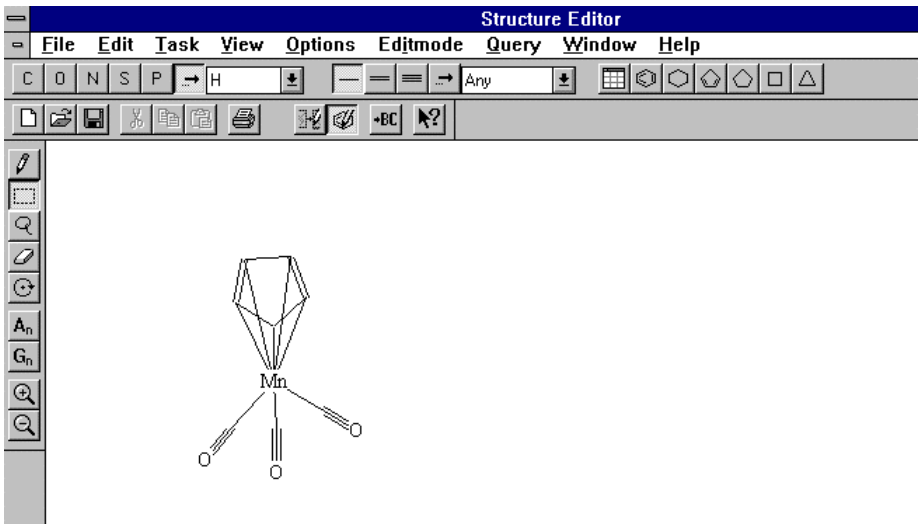
What you will learn:

- ***Combining substance queries with product registry numbers***
- ***Cutting down the reaction display using the Fact Editor and the Hit only option in the display***

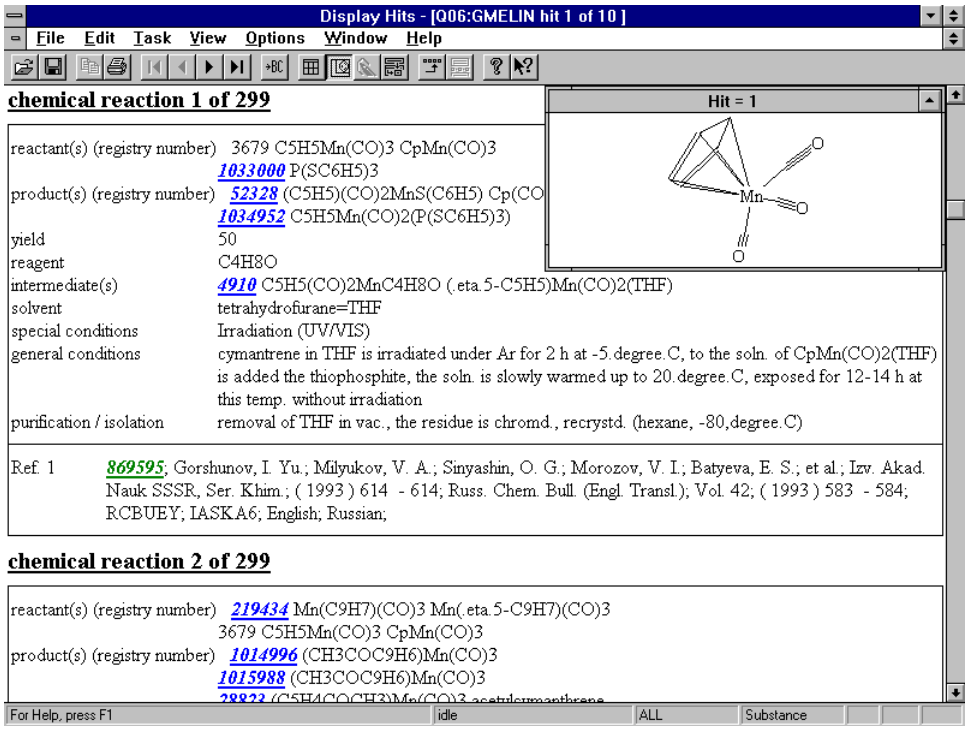
We want to search for preparationsof Cymantrene (cyclopentadienyl manganese tricarbonyl)

8.1 Structure Query

Draw the manganese complex in the structure editor:



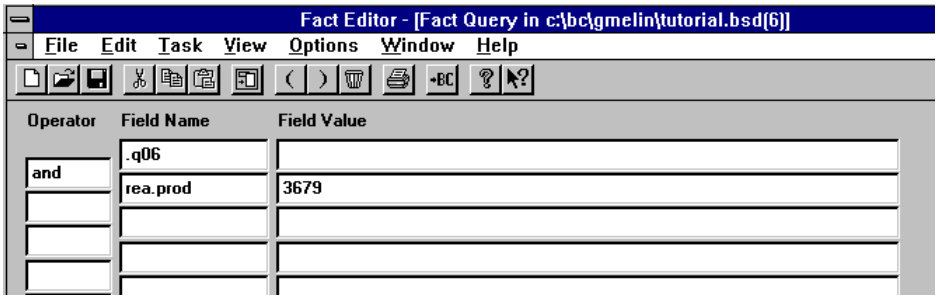
Transfer the structure to the Commander and perform the search. You will receive 10 hits. When you take a look at them in the short display you will notice that the first hit is the compound we have been looking for; the others are multi-component or isotope labelled compounds. Double click the first hit and move down the record to the reactions. You will see 299 reactions:



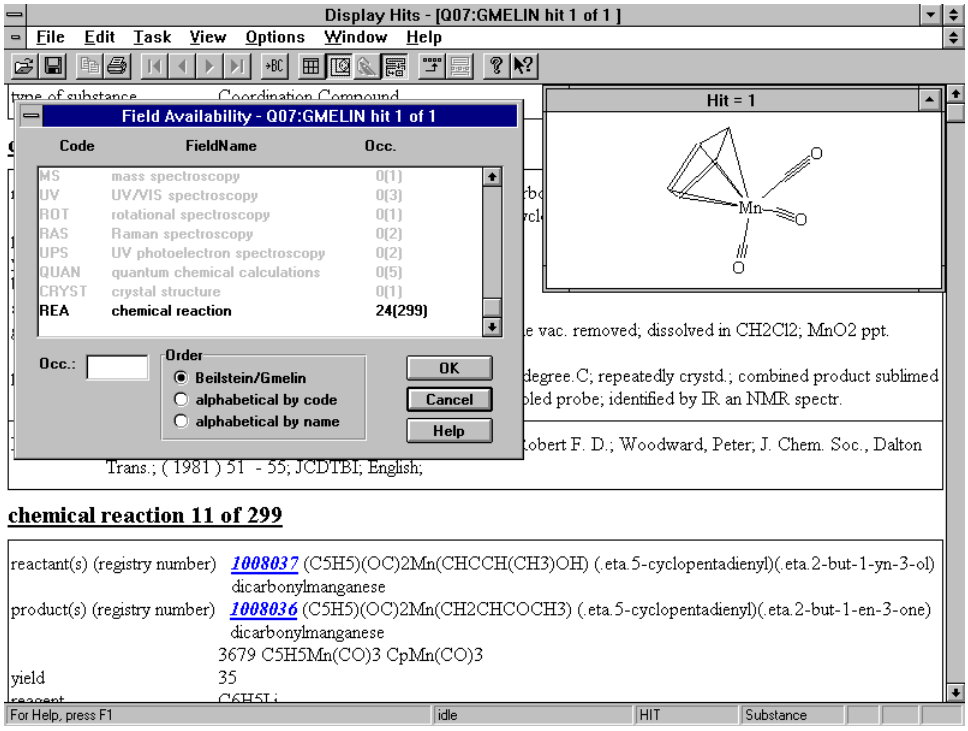
In most of them Cymantrene acts as a reactant, but at the moment we are only interested in ways of preparing Cymantrene. Therefore we will employ a trick to cut down the display. If you combine the hitset you just retrieved with a query for the registry number of the desired product (i. e. Cymantrene) and use the Hit only option in the display afterwards you will see only those reactions in which Cymantrene is the product of the reaction. Note now the registry number of Cymantrene, which is 3679.

8.2 Combination of the Former Hitset with a Fact Query for a Product Registry Number

In the next step we will formulate a fact query in the fact editor. Load the number of the former hitset (the original structure query) or type in the number directly, if you remember it. Then combine this query with the query for the product registry number. This entry can be found in the reaction section of the datastructure:



Start the search and you will receive one hit, which is of course the first hit of the former hitset. Switch to the "Hit only" display and call up the Field availability list. You will notice that now only 24 reactions are shown, all the others are hidden. If you scroll through those reaction data you will see that these reactions are indeed only preparations of Cymantrene:



By this procedure you can save much time since you do not have to scroll through all reaction data of the product that interests you.

The same procedure can be applied to look for reactions of a special educt by combining the original query with a reactant registry number. Then you will get a display of only those reactions in which the title compound acts as reactant when you employ the Hit only view.

After having finished this last chapter you will have learned about all important features of the Gmelin database under CrossFire. It is now up to you to explore the data contained there to help you with your research.

9 Index